



US009284571B2

(12) **United States Patent**  
**Damude et al.**

(10) **Patent No.:** **US 9,284,571 B2**  
(45) **Date of Patent:** **Mar. 15, 2016**

(54) **USE OF A SEED SPECIFIC PROMOTER TO DRIVE ODP1 EXPRESSION IN CRUCIFEROUS OILSEED PLANTS TO INCREASE OIL CONTENT WHILE MAINTAINING NORMAL GERMINATION**

(71) Applicant: **E.I. du Pont de Nemours and Company**, Wilmington, DE (US)

(72) Inventors: **Howard Glenn Damude**, Hockessin, DE (US); **John D. Everard**, Wilmington, DE (US); **Knut Meyer**, Wilmington, DE (US); **Kevin G. Ripp**, Wilmington, DE (US); **Kevin L. Stecca**, Bear, DE (US)

(73) Assignee: **E I DU PONT DE NEMOURS AND COMPANY**, Wilmington, DE (US)

(\* ) Notice: Subject to any disclaimer, the term of this patent is extended or adjusted under 35 U.S.C. 154(b) by 379 days.

(21) Appl. No.: **13/769,539**

(22) Filed: **Feb. 18, 2013**

(65) **Prior Publication Data**

US 2013/0185825 A1 Jul. 18, 2013  
US 2013/0347142 A9 Dec. 26, 2013

**Related U.S. Application Data**

(62) Division of application No. 12/752,175, filed on Apr. 1, 2010, now Pat. No. 8,404,926.

(60) Provisional application No. 61/165,548, filed on Apr. 1, 2009.

(51) **Int. Cl.**  
**C12N 15/82** (2006.01)

(52) **U.S. Cl.**  
CPC ..... **C12N 15/8247** (2013.01); **C12N 15/8201** (2013.01)

(58) **Field of Classification Search**  
None  
See application file for complete search history.

(56) **References Cited**

**U.S. PATENT DOCUMENTS**

5,107,065 A 4/1992 Shewmaker et al.  
5,968,809 A 10/1999 Knutzon et al.  
6,072,050 A 6/2000 Bowen et al.  
6,512,165 B1 1/2003 Ross et al.  
6,555,673 B1 4/2003 Bowen et al.  
7,157,621 B2 1/2007 Allen et al.  
2003/0135889 A1 7/2003 Ross et al.  
2003/0226166 A1 12/2003 Falco et al.

**FOREIGN PATENT DOCUMENTS**

WO WO 9967405 A2 12/1999  
WO WO 0028058 A2 5/2000  
WO WO 0200904 A2 1/2002  
WO WO 0208269 A2 1/2002  
WO WO 03001902 A2 1/2003  
WO WO 2004071467 A2 8/2004

WO WO 2005075655 A2 8/2005  
WO WO 2006000732 A1 1/2006  
WO WO 2007061845 A2 5/2007

**OTHER PUBLICATIONS**

Bowie et al, *Science* 247:1306-1310, 1990.\*  
McConnell et al, *Nature* 411 (6838):709-713, 2001.\*  
Marsch-Martinez et al (2006, *Plant Mol. Biol.* 62:825-843).\*  
Boutillier et al (2002, *The Plant Cell* 14:1737-1749).\*  
Benfey et al (1990, *Science* 250:959-966)\*  
Benfey et al (1989, *EMBO J.* 8(8):2195-2202.\*  
U.S. Appl. No. 61/578,903, filed Dec. 22, 2011, Damude.  
Baud, S., et al. "A Spatiotemporal Analysis of Enzymatic Activities Associated With Carbon Metabolism in Wild-Type and Mutant Embryos of Arabidopsis Using In Situ Histochemistry." *The Plant Journal*, vol. 46:155-169, 2006.  
Baud, S., et al., "Structure and Expression Profile of the Sucrose Synthase Multigene Family in Arabidopsis," *Journal of Experimental Botany*, vol. 55(396):397-409, 2004.  
Beachy, R.N., et al., "Accumulation and Assembly of Soybean B-Conglycinin in Seeds of Transformed Petunia Plants," *The Embo Journal*, vol. 4(12):3047-3053, 1985.  
Becker, D.M., et al., "A CDNA Encoding a Human CCAAT-Binding Protein Cloned by Functional Complementation in Yeast," *Proc. Natl. Acad. Sci. USA*, vol. 88:1968-1972, 1991.  
Boutillier, K., et al., "Ectopic Expression of BABY BOOM Triggers a Conversion from Vegetative to Embryonic Growth," *The Plant Cell* 14: 1737-1749, Aug. 2002.  
Bowie, J.U., et al., "Deciphering the Message in Protein Sequences: Tolerance to Amino Acid Substitutions," *Science* 247:1306-1310, 1990.  
Cernac, A., et al., "Wrinkled I Encodes an AP2/EREB Domain Protein Involved in the Control of Storage Compound Biosynthesis in Arabidopsis," *The Plant Journal*, 40:575-585. 2004.  
Drews, G.N., et al., "Negative Regulation of the Arabidopsis Homeotic Gene *Agamous* by the *Apetala2* Product," *Cell*, vol. 65(6):991-1002, 1991.  
Edwards, D., et al., "Multiple Genes Encoding the Conserved CCAAT-Box Transcription Factor Complex are Expressed in Arabidopsis," *Plant Physiol.*, vol. 117:1015-1022, 1998.  
Goldberg, R.B., et al., "Regulation of Gene Expression During Plant Embryogenesis," *Cell*, vol. 56(2):149-160, 1989.  
National Center for Biotechnology Information General Identifier No. 32364685, Accession No. AAP80382, Aug. 23, 2004, A.Cernac et al., "Wrinkled1 [Arabidopsis Thaliana]," *Biochemistry and Molecular Biology*, MSU, East Lansing, Michigan, USA.  
Irish, V.F., et al., "Function of the *Apetala-1* Gene During Arabidopsis Floral Development," *The Plant Cell*, vol. 2:741-753, 1990.  
Jackson, S.M., et al., "NF-Y Has a Novel Role in Sterol-Dependent Transcription of Two Cholesterogenic Genes," *The Journal of Biological Chemistry*, vol. 270(37):21445-21448, 1995.  
Jofuku, D., et al., "Control of Arabidopsis Flower and Seed Development by the Homeotic Gene *Apetala2*," *The Plant Cell*, vol. 6:1211-1225, 1994.  
Kagaya, Y., et al., "The promoter from the rice nuclear gene encoding chloroplast aldolase confers mesophyll-specific and light-regulated expression in transgenic tobacco," *Mol. Gen. Genet.*, 248: 668-674 1995.  
Li, X, et al., "Evolutionary Variation of the CCAAT-Binding Transcription Factor NF-Y," *Nucleic Acids Research*, vol. 20(5):1087-1091, 1991.  
Li, Y., et al., "Oil Content of Arabidopsis Seeds: The Influence of Seed Anatomy, Light and Plant-to-Plant Variation," *Elsevier Phytochemistry*, vol. 67:904-915, 2006.

(Continued)

*Primary Examiner* — Stuart F Baum

(57) **ABSTRACT**

A recombinant DNA construct comprising a polynucleotide encoding an ODP1 polypeptide operably linked to a sucrose synthase 2 promoter where this construct can be used to increase oil content in the seeds of a cruciferous oilseed plant while maintaining normal germination is disclosed. A method for increasing oil content in the seeds of a cruciferous oilseed plant while maintaining normal germination using this construct is also disclosed.

**23 Claims, 4 Drawing Sheets**

(56)

**References Cited**

## OTHER PUBLICATIONS

Lopez, J.M. et al., "Sterol Regulation of Acetyl Coenzyme A Carboxylase: A Mechanism for Coordinate Control of Cellular Lipid," *Proc. Natl. Acad. Sci. USA*, vol. 93:1049-1053, 1996.

Lotan, T., et al., "Arabidopsis Leafy Cotyledon1 Is Sufficient to Induce Embryo Development in Vegetative Cells," *Cell*, vol. 93:1195-1205, 1998.

Marsh-Martinez, N., et al., BOLITA, an Arabidopsis AP2/ERF-like transcription factor that affects cell expansion and proliferation/differentiation pathways, *Plant Mol. Biol.* 62: 825-843, 2006.

McConnell, J.R., et al., "Role of Phabulosa and Phavoluta in determining radial patterning in shoots," *Nature* 411 (6838):709-713, 2001.

McKnight, S.L., et al., "Is CCAAT/Enhancer-Binding Protein a Central Regulator of Energy Metabolism?," *Genes Dev.* vol. 3:2021-2024, 1989.

Okamuro, J.K., et al., "The Ap2 Domain of Apetala2 Defines a Large New Family of DNA Binding Proteins in Arabidopsis," *Proc. Natl. Acad. Sci. USA*, vol. 94:7076-7081, 1997.

Roder, K., et al., "NF-Y Binds to the Inverted CCAAT Box, An Essential Element for C Amp-dependent Regulation of the Rat Fatty Acid Synthase (FAS) Gene," *Gene*, vol. 184:21-26, 1997.

Ruuska, S.A., et al., "Contrapuntal Networks of Gene Expression During Arabidopsis Seed Filling," *The Plant Cell*, vol. 14:1191-1206, 2002.

Sinha, S., et al., "Recombinant Rat CBF-C, The Third Subunit of CBF/NFY, Allows Formation of a Protein-DNA Complex With CBF-A and CBF-B and With Yeast HAP2 and HAP3," *Proc. Natl. Acad. Sci. USA*, vol. 92:1624-1628, 1995.

Ohme-Takagi, M., et al., "Ethylene-Inducible DNA Binding Proteins That Interact With an Ethylene-Responsive Element," *The Plant Cell*, vol. 7:173-182, 1995.

Johan Ericsson et al., "Synergistic Binding of Sterol Regulatory Element-Binding Protein and NF-Y to the Farnesyl Diphosphate Synthase Promoter is Critical for Sterol-Regulated Expression of the Gene", *The Journal of Biological Chemistry*, vol. 271(40):24359-24364, 1996.

Juan Gabriel Angeles-Nunez et al., Regulation of AtSUS2 and AtSUS3 by glucose and the transcription factor LEC2 in different tissues and at different stages of Arabidopsis seed development, *Plant Mol Biol* (2012) 78:377-392.

\* cited by examiner





FIG. 1C

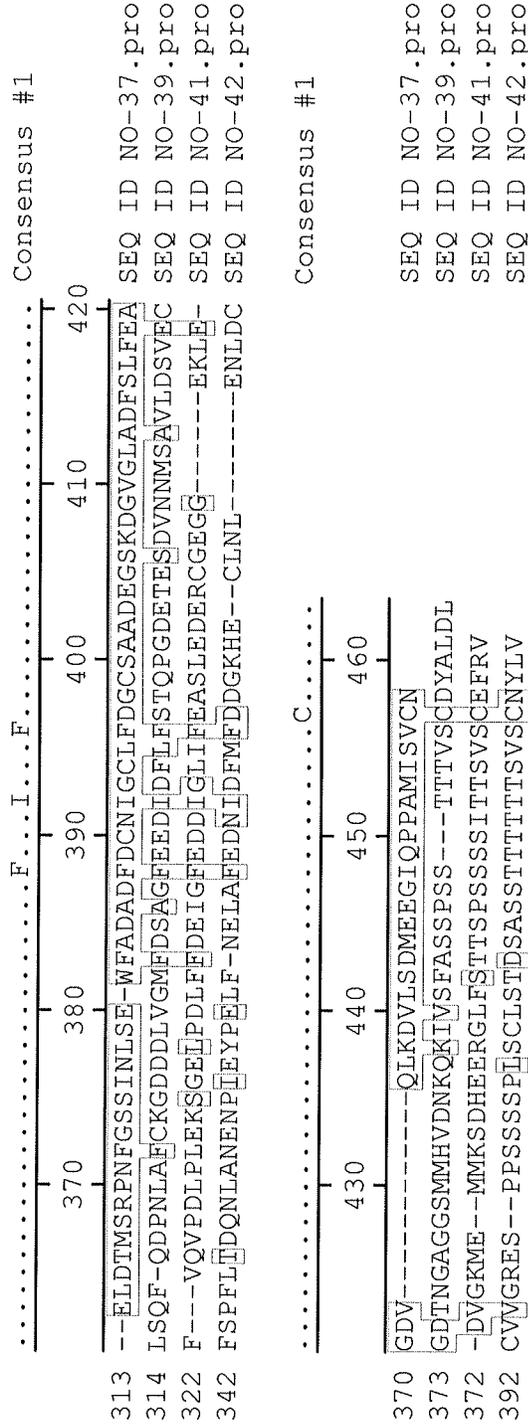


FIG. 2

		Percent Identity				
		1	2	3	4	
1		████████	41.8	43.5	41.8	1
2		79.5	████████	46.5	45.4	2
3		78.0	59.2	████████	48.4	3
4		85.0	72.3	59.1	████████	4
		1	2	3	4	
	Divergence					
						SEQ ID NO-37.pro
						SEQ ID NO-39.pro
						SEQ ID NO-41.pro
						SEQ ID NO-42.pro

1

**USE OF A SEED SPECIFIC PROMOTER TO  
DRIVE ODP1 EXPRESSION IN  
CRUCIFEROUS OILSEED PLANTS TO  
INCREASE OIL CONTENT WHILE  
MAINTAINING NORMAL GERMINATION**

CROSS REFERENCE TO RELATED  
APPLICATIONS

This application is a divisional of U.S. patent application Ser. No. 12/752,175, filed Apr. 1, 2010, now U.S. Pat. No. 8,404,926, issued on Mar. 26, 2013, which claims the benefit of U.S. Provisional Application No. 61/165,548, filed Apr. 1, 2009, the entire content of each is herein incorporated by reference.

REFERENCE TO SEQUENCE LISTING  
SUBMITTED ELECTRONICALLY

The official copy of the sequence listing is submitted electronically via EFS-Web as an ASCII formatted sequence listing with a file named 429238seqlist.txt, created on Feb. 12, 2013, and having a size of 604 KB and is filed concurrently with the specification. The sequence listing contained in this ASCII formatted document is part of the specification and is herein incorporated by reference in its entirety

FIELD OF THE INVENTION

This invention is in the field of biotechnology, in particular, this pertains to increasing oil content while maintaining normal germination in a cruciferous oilseed plant using a seed specific promoter to drive expression of ODP1.

BACKGROUND OF THE INVENTION

Plant lipids have a variety of industrial and nutritional uses and are central to plant membrane function and climatic adaptation. These lipids represent a vast array of chemical structures, and these structures determine the physiological and industrial properties of the lipid. Many of these structures result either directly or indirectly from metabolic processes that alter the degree of unsaturation of the lipid. Different metabolic regimes in different plants produce these altered lipids, and either domestication of exotic plant species or modification of agronomically adapted species is usually required to produce economically large amounts of the desired lipid.

There are serious limitations to using mutagenesis to alter fatty acid composition and content. Screens will rarely uncover mutations that a) result in a dominant ("gain-of-function") phenotype, b) are in genes that are essential for plant growth, and c) are in an enzyme that is not rate-limiting and that is encoded by more than one gene. In cases where desired phenotypes are available in mutant corn lines, their introgression into elite lines by traditional breeding techniques is slow and expensive, since the desired oil compositions are likely the result of several recessive genes.

Recent molecular and cellular biology techniques offer the potential for overcoming some of the limitations of the mutagenesis approach, including the need for extensive breeding. Some of the particularly useful technologies are seed-specific expression of foreign genes in transgenic plants (see Goldberg et al (1989) *Cell* 56:149-160), and the use of antisense RNA to inhibit plant target genes in a dominant and tissue-specific manner (see van der Krol et al (1988) *Gene* 72:45-50). Other advances include the transfer of foreign

2

genes into elite commercial varieties of commercial oilcrops, such as soybean (Chee et al (1989) *Plant Physiol.* 91:1212-1218; Christou et al (1989) *Proc. Natl. Acad. Sci. U.S.A.* 86:7500-7504; Hinchee et al (1988) *Bio/Technology* 6:915-922; EPO publication 0 301 749 A2), rapeseed (De Block et al (1989) *Plant Physiol.* 91:694-701), and sunflower (Everett et al (1987) *Bio/Technology* 5:1201-1204), and the use of genes as restriction fragment length polymorphism (RFLP) markers in a breeding program, which makes introgression of recessive traits into elite lines rapid and less expensive (Tank-  
sley et al (1989) *Bio/Technology* 7:257-264). However, application of each of these technologies requires identification and isolation of commercially-important genes.

Transcription factors generally bind DNA in a sequence-specific manner and either activate or repress transcription initiation. The specific mechanisms of these interactions remain to be fully elucidated. At least three types of separate domains have been identified within transcription factors. One is necessary for sequence-specific DNA recognition, one for the activation/repression of transcriptional initiation, and one for the formation of protein-protein interactions (such as dimerization). Studies indicate that many plant transcription factors can be grouped into distinct classes based on their conserved DNA binding domains (Katagiri F and Chua N H, 1992, *Trends Genet.* 8:22-27; Menkens A E, Schindler U and Cashmore A R, 1995, *Trends in Biochem Sci.* 13:506-510; Martin C and Paz-Ares J, 1997, *Trends Genet.* 13:67-73). Each member of these families interacts and binds with distinct DNA sequence motifs that are often found in multiple gene promoters controlled by different regulatory signals.

Several transcription factor families have been identified in plants. For example, nucleotide sequences encoding the following transcription factors families have been identified: Alfin-like, AP2 (APETALA2) and EREBPs (ethylene-responsive element binding proteins), ARF, AUX/IAA, bHLH, bZIP, C2C2 (Zn), C2C2 (Co-like), C2C2 (Dof), C2C2 (GATA), C2C2 (YABBY), C2H2 (Zn), C3H-type, CCAAT, CCAAT HAP3, CCAAT HAP5, CPP (Zn), DRAP1, E2F/DP, GARP, GRAS, HMG-BOX, HOMEO BOX, HSF, Jumanji, LFY, LIM, MADS Box, MYB, NAC, NIN-like, Polycomb-like, RAV-like, SBP, TCP, TFIID, Transfactor, Trihelix, TUBBY, and WRKY.

WO 2005/075655 published on Aug. 18, 2005 describes an AP2 domain transcription factor ODP2 (ovule development protein 2) and methods of U.S. Pat. No. 7,157,621 which issued on Jan. 2, 2007, describes the alteration of oil traits in plants through controlled expression of selected genes in plants.

The AP2/ERF family of proteins is a plant-specific class of putative transcription factors that have been shown to regulate a wide-variety of developmental processes and are characterized by the presence of an AP2/ERF DNA binding domain. The AP2/ERF proteins have been subdivided into two distinct subfamilies based on whether they contain one (ERF subfamily) or two (AP2 subfamily) DNA binding domains.

Specifically, AP2 (APETALA2) and EREBPs (ethylene-responsive element binding proteins) are the prototypic members of a family of transcription factors unique to plants, whose distinguishing characteristic is that they contain the so-called AP2 DNA-binding domain. AP2/EREBP genes form a large multigene family, and they play a variety of roles throughout the plant life cycle. AP2/EREBP genes are key regulators of several developmental processes, including floral organ identity determination and leaf epidermal cell identity. In *Arabidopsis thaliana*, the homeotic gene APETALA2 (AP2) has been shown to control three salient processes during development: (1) the specification of flower organ identity

throughout floral organogenesis (Jofuku et al., *Plant Cell* 6:1211-1225, 1994); (2) establishment of flower meristem identity (Irish and Sussex, *Plant Cell* 2:8:741-753, 1990); and (3) the temporal and spatial regulation of flower homeotic gene activity (Drews et al., *Cell* 65:6:991-1002, 1991). DNA sequence analysis suggests that AP2 encodes a theoretical polypeptide of 432 aa, with a distinct 68 aa repeated motif termed the AP2 domain. This domain has been shown to be essential for AP2 functions and contains within the 68 aa, an eighteen amino acid core region that is predicted to form an amphipathic  $\alpha$ -helix (Jofuku et al., *Plant Cell* 6:1211-1225, 1994). AP2-like domain-containing transcription factors have been also identified in both *Arabidopsis thaliana* (Okamuro et al., *Proc. Natl. Acad. Sci. USA* 94:7076-7081, 1997) and in tobacco with the identification of the ethylene responsive element binding proteins (EREBPs) (Ohme-Takagi and Shinshi, *Plant Cell* 7:2:173-182, 1995). In *Arabidopsis*, these RAP2 (related to AP2) genes encode two distinct subfamilies of AP2 domain-containing proteins designated AP2-like and EREBP-like (Okamuro et al., *Proc. Natl. Acad. Sci. USA* 94:7076-7081, 1997). In vitro DNA binding has not been shown to date using the RAP2 proteins. Based upon the presence of two highly conserved motifs YRG and RAYD within the AP2 domain, it has been proposed that binding DNA binding occurs in a manner similar to that of AP2 proteins.

As was noted above, regulation of transcription of most eukaryotic genes is coordinated through sequence-specific binding of proteins to the promoter region located upstream of the gene. Many of these protein-binding sequences have been conserved during evolution and are found in a wide variety of organisms. One such feature is the "CCAAT" sequence element (Edwards et al, 1998, *Plant Physiol.* 117: 1015-1022). CCAAT boxes are a feature of gene promoters in many eukaryotes including several plant gene promoters.

HAP proteins constitute a large family of transcription factors first identified in yeast. They combine to form a heteromeric protein complex that activates transcription by binding to CCAAT boxes in eukaryotic promoters. The orthologous HAP proteins display a high degree of evolutionary conservation in their functional domains in all species studied to date (Li et al., 1991, *Nucleic Acids Res.* 20:1087-1091).

WO 00/28058 published on May 18, 2000 describes HAP3-type CCAAT-box binding transcriptional activator polynucleotides and polypeptides, especially, the leafy cotyledon 1 transcriptional activator (LEC1) polynucleotides and polypeptides.

WO 99/67405 describes leafy cotyledon1 genes and their uses.

The human, murine and plant homologues of CCAAT-binding proteins have been isolated and characterized based on their sequence similarity with their yeast counterparts (Li et al., 1991, *Nucleic Acids Res.* 20:1087-1091). This high degree of sequence homology translates remarkably into functional interchangeability among orthologue proteins of different species (Sinha et al, 1995, *Proc. Natl. Acad. Sci. USA* 92:1624-1628). Unlike yeast, multiple forms of each HAP homologue have been identified in plants (Edwards et al, 1998, *Plant Physiol.* 117:1015-1022).

Molecular and genetic analysis revealed HAP members to be involved in the control of diverse and critical biological processes ranging from development and cell cycle regulation to metabolic control and homeostasis (Lotan et al, 1998, *Cell* 93:1195-1205; Lopez et al, 1996, *Proc. Natl. Acad. Sci. USA* 93:1049-1053). In yeast, HAPs are involved in the transcriptional control of metabolic processes such as the regu-

lation of catabolic derepression of cycl and other genes involved in respiration (Becker et al., 1991, *Proc. Natl. Acad. Sci. USA* 88:1968-1972).

In mammalian systems, several reports describe HAPs as direct or indirect regulators of several important genes involved in lipid biosynthesis such as fatty acid synthase (Roder et al, 1997, *Gene* 184:21-26), farnesyl diphosphate (FPP) synthase (Jackson et al, 1995, *J. Biol. Chem.* 270: 21445-21448; Ericsson et al, 1996, *J. Biol. Chem.* 271:24359-24364), glycerol-3-phosphate acyltransferase (GPA, Jackson et al, 1997), acetyl-CoA carboxylase (ACC, Lopez et al, 1996, *Proc. Natl. Acad. Sci. USA* 93:1049-1053) and 3-hydroxy-3-methylglutaryl-coenzyme A (HMG-CoA) synthase (Jackson et al, 1995, *J. Biol. Chem.* 270:21445-21448), among others.

In addition, other CCAAT-binding transcription factors have also been reported to be involved in different aspects of the control of lipid biosynthesis and adipocyte growth and differentiation in mammalian systems (see McKnight et al, 1989).

It appears that the currently available evidence to date points to a family of proteins of the CCAAT-binding transcription factors as important modulators of metabolism and lipid biosynthesis in mammalian systems. Such a determination has not been made for plant systems.

Other polypeptides that influence ovule and embryo development and stimulate cell growth, such as, Lec1, Kn1, WUSCHEL, Zwillie and Aintegumeta (ANT) allow for increased transformation efficiencies when expressed in plants. See, for example, U.S. Application No. 2003/0135889, herein incorporated by reference. In fact, a maize Lec1 homologue of the *Arabidopsis* embryogenesis controlling gene AtLEC1, has been shown to increase oil content and transformation efficiencies in plants. See, for example, WO 03001902 and U.S. Pat. No. 6,512,165.

The putative AP2/EREBP transcription factor WRINKLED1 (WRI1) is involved in the regulation of seed storage metabolism in *Arabidopsis* (Cermac and Benning, 2004, *Plant J.* 40:575-585). Expression of the WRI1<sup>TM</sup> cDNA under the control of the CaMV 35S promoter led to increased seed oil content. Oil-accumulating seedlings, however, showed aberrant development consistent with a prolonged embryonic state. Nucleic acid molecules encoding WRINKLED1-LIKE polypeptides and methods of use are also described in International Publication No. WO 2006/00732 A2.

Because transcription factors regulate transcription and orchestrate gene expression in plants and other organisms, control of transcription factor gene expression provides a powerful means for altering plant phenotype. The transformation of plants with transcription factors, however, can result in aberrant development based on the overexpression and/or ectopic expression of the transcription factor. In the current invention, it has been found that use of a seed specific promoter, such as SUS2 from *Arabidopsis*, can drive expression of an ODP1 gene thereby increasing oil content in the seeds of a cruciferous oilseed plant without negatively affecting germination and seedling establishment.

#### SUMMARY OF THE INVENTION

In a first embodiment, the present invention concerns a recombinant DNA construct comprising a polynucleotide encoding an ODP1 polypeptide operably linked to a sucrose synthase 2 promoter wherein said construct increases oil content in the seeds of a cruciferous oilseed plant while maintaining normal germination and further wherein the amino

acid sequence of said ODP1 polypeptide has at least 80%, at least 90%, at least 95% or 100% sequence identity to a sequence selected from the group consisting of SEQ ID NO:37, SEQ ID NO:39, and SEQ ID NO:41.

In another embodiment, the present invention concerns a recombinant construct comprising a sucrose synthase 2 promoter which comprises: (a) the nucleotide sequence of SEQ ID NO:43, SEQ ID NO:72 or SEQ ID NO:73, or (b) a nucleotide sequence comprising a functional fragment of the nucleotide sequence of SEQ ID NO:43, SEQ ID NO:72 or SEQ ID NO:73.

In another embodiment, the present invention concerns a transgenic cruciferous oilseed plant comprising in its genome the recombinant DNA construct of the invention. Also included are transgenic seeds obtained from such transgenic cruciferous oilseed plants, wherein the transgenic seed comprises in its genome the recombinant DNA construct of the invention.

In another embodiment, the present invention concerns a method for producing a transgenic cruciferous oilseed plant comprising transforming a cruciferous oilseed plant cell with the recombinant construct of the invention and regenerating a transgenic plant from the transformed plant cell, wherein the transgenic cruciferous oilseed plant comprises in its genome the recombinant DNA construct of the invention.

In another embodiment, the present invention concerns a method for increasing oil content in seeds of a transgenic cruciferous oilseed plant while maintaining normal germination, said method comprising:

- (a) transforming a cruciferous oilseed plant cell with a recombinant DNA construct comprising a polynucleotide encoding an ODP1 polypeptide, wherein the amino acid sequence of said ODP1 polypeptide has at least 80%, at least 90% or at least 95% sequence identity with a sequence selected from the group consisting of SEQ ID NO:37, SEQ ID NO:39, and SEQ ID NO:41, said sequence being operably linked to a seed specific promoter;
- (b) regenerating a transgenic cruciferous oilseed plant from the transformed cell of step (a), wherein said plant comprises the recombinant DNA construct;
- (c) obtaining a transgenic progeny plant derived from the transgenic cruciferous oilseed plant of step (b), wherein the transgenic progeny plant comprises in its genome the recombinant DNA construct;
- (d) assaying the transgenic progeny plant obtained from step (c) for oil level and germination; and
- (e) selecting those transgenic progeny plants having seeds with an increased level of oil and normal germination when compared to seeds obtained from a control cruciferous oilseed plant, wherein said control plant does not comprise the recombinant DNA construct.

In another embodiment, the present invention concerns a method of the invention wherein the ODP1 polypeptide is a maize ODP1 polypeptide and, more specifically, the amino acid sequence of the ODP1 polypeptide comprises the sequence of SEQ ID NO:37. In addition, the seed specific promoter can be a sucrose synthase 2 promoter and, more specifically, the nucleotide sequence of sucrose synthase 2 promoter comprises (a) the nucleotide sequence of SEQ ID NO:43, SEQ ID NO:72 or SEQ ID NO:73; or (b) a nucleotide sequence comprising a functional fragment of the nucleotide sequence of SEQ ID NO:43, SEQ ID NO:72 or SEQ ID NO:73.

In another embodiment, the present invention concerns oil or by-products obtained from transgenic seed of the invention.

In another embodiment, the cruciferous oilseed plant or seed of any of the compositions or methods of the present invention can be canola or *Arabidopsis* or other plant species including but not limited to the following: *Barbarea vulgaris*, *Brassica campestris*, *Brassica carinata*, *Brassica elongata*, *Brassica fruticulosa*, *Brassica hirta*, *Brassica juncea*, *Brassica napus*, *Brassica narinosa*, *Brassica nigra*, *Brassica oleracea*, *Brassica perviridis*, *Brassica rapa*, *Brassica rupestris*, *Brassica septiceps*, *Brassica tournefortii*, *Brassica verna*, *Camelina sativa*, *Crambe abyssinica*, *Lepidium campestre*, *Raphanus sativus*, *Sinapis alba*.

#### BRIEF DESCRIPTION OF THE DRAWINGS AND SEQUENCE LISTING

The invention can be more fully understood from the following detailed description and the accompanying drawings and Sequence Listing which form a part of this application.

FIG. 1A-1C show a multiple alignment of the ODP1 polypeptides of *Zea mays* (SEQ ID NO:37), *Glycine max* (SEQ ID NO:39), *Momordica charantia* (SEQ ID NO:41), and the WRINKLED1 gene from *Arabidopsis thaliana* (SEQ ID NO:42; NCBI GI NO. 32364685). The multiple alignment was assembled using the Clustal V method of alignment with the default parameters. Residues that match SEQ ID NO:37 exactly are enclosed in a box. Above the alignment is shown a consensus sequence (SEQ ID NO:74). A residue is shown in the consensus sequence when all residues at that position are identical.

FIG. 2 shows the percent sequence identity and divergence for each pair of polypeptides from the multiple alignment of FIG. 1A-1C.

SEQ ID NO:1 is the nucleotide sequence of vector pKS121/BS.

SEQ ID NO:2 is the nucleotide sequence of vector pDsRedxKS121/BS.

SEQ ID NO:3 is the nucleotide sequence of vector pKS332.

SEQ ID NO:4 is the nucleotide sequence of PCR primer MWG345.

SEQ ID NO:5 is the nucleotide sequence of PCR primer MWG346.

SEQ ID NO:6 is the nucleotide sequence of vector pKS336.

SEQ ID NO:7 is the nucleotide sequence of the T-DNA of the plant transformation vector pZBL120xKS336.

SEQ ID NO:8 is the nucleotide sequence of PCR primer MWG339.

SEQ ID NO:9 is the nucleotide sequence of PCR primer MWG340.

SEQ ID NO:10 is the nucleotide sequence of vector pKS333.

SEQ ID NO:11 is the nucleotide sequence of the T-DNA of the plant transformation vector pZBL120xKS333.

SEQ ID NO:12 is the nucleotide sequence of PCR primer MWG341.

SEQ ID NO:13 is the nucleotide sequence of PCR primer MWG342.

SEQ ID NO:14 is the nucleotide sequence of vector pKS334.

SEQ ID NO:15 is the nucleotide sequence of the T-DNA of the plant transformation vector pZBL120xKS334.

SEQ ID NO:16 is the nucleotide sequence of vector pKR132.

SEQ ID NO:17 is the nucleotide sequence of vector pKR627.

SEQ ID NO:18 is the nucleotide sequence of vector KS294.

SEQ ID NO:19 is the nucleotide sequence of vector pKR1142.

SEQ ID NO:20 is the nucleotide sequence of vector pKR1141.

SEQ ID NO:21 is the nucleotide sequence of PCR primer SuSy-5.

SEQ ID NO:22 is the nucleotide sequence of PCR primer SuSy-3.

SEQ ID NO:23 is the nucleotide sequence of vector pLF122.

SEQ ID NO:24 is the nucleotide sequence of vector pKR1155.

SEQ ID NO:25 is the nucleotide sequence of vector pKR1158.

SEQ ID NO:26 is the nucleotide sequence of vector pKR1167.

SEQ ID NO:27 is the nucleotide sequence of vector pKR92.

SEQ ID NO:28 is the nucleotide sequence of vector pKR1223.

SEQ ID NO:29 is the nucleotide sequence of vector pKR268.

SEQ ID NO:30 is the nucleotide sequence of vector pKR1143.

SEQ ID NO:31 is the nucleotide sequence of vector pKR1147.

SEQ ID NO:32 is the nucleotide sequence of vector pKR1220.

SEQ ID NO:33 is the nucleotide sequence of vector pKR1144.

SEQ ID NO:34 is the nucleotide sequence of vector pKR1149.

SEQ ID NO:35 is the nucleotide sequence of vector pKR1221.

SEQ ID NO:36 is the nucleotide sequence of the maize ODP1 coding region from cDNA clone cde1c.pk003.o22.

SEQ ID NO:37 is the amino acid sequence of the maize ODP1 encoded by SEQ ID NO:36. SEQ ID NO:37 is identical to SEQ ID NO:320 in U.S. Pat. No. 7,157,621.

SEQ ID NO:38 is the nucleotide sequence of the soybean ODP1 coding region from cDNA clone se3.pk0003.f5.

SEQ ID NO:39 is the amino acid sequence of the soybean ODP1 encoded by SEQ ID NO:38. SEQ ID NO:39 is identical to SEQ ID NO:481 in U.S. Pat. No. 7,157,621.

SEQ ID NO:40 is the nucleotide sequence of the *Momordica charantia* ODP1 coding region from cDNA clone fds1n.pk015.115.

SEQ ID NO:41 is the amino acid sequence of the *Momordica charantia* ODP1 encoded by SEQ ID NO:40. SEQ ID NO:41 is identical to SEQ ID NO:477 in U.S. Pat. No. 7,157,621.

SEQ ID NO:42 is the amino acid sequence of WRINKLED1 (WR11) from *Arabidopsis thaliana* and corresponds to NCBI GI NO. 32364685.

SEQ ID NO:43 is the nucleotide sequence of the sucrose synthase 2 (SUS2) promoter from *Arabidopsis thaliana* that is present in vector pKR1223.

SEQ ID NO:44 is the nucleotide sequence of the canola SUS2 homolog.

SEQ ID NO:45 is the amino acid sequence of the canola SUS2 homolog encoded by SEQ ID NO:44.

SEQ ID NO:46 is the nucleotide sequence of primer a.

SEQ ID NO:47 is the nucleotide sequence of primer b.

SEQ ID NO:48 is the nucleotide sequence of primer c.

SEQ ID NO:49 is the nucleotide sequence of primer d.

SEQ ID NO:50 is the nucleotide sequence of "PvuII rapa cons", a genomic sequence of canola variety NS1822BC that was generated with primers a and b.

SEQ ID NO:51 is the nucleotide sequence of "1,6 DraI gene cons", a genomic sequence of canola variety NS1822BC that was generated with primers c and d.

SEQ ID NO:52 is the nucleotide sequence of primer SA188.

SEQ ID NO:53 is the nucleotide sequence of primer SA189.

SEQ ID NO:54 is the nucleotide sequence of primer SA190.

SEQ ID NO:55 is the nucleotide sequence of primer SA191.

SEQ ID NO:56 is the nucleotide sequence of "BN SUS2 prom1/PCR blunt", which is derived from 1,6 DraI gene cons (SEQ ID NO:51).

SEQ ID NO:57 is the nucleotide sequence of "BN SUS2 prom2/PCR blunt", which is derived from PvuII rapa cons (SEQ ID NO:50).

SEQ ID NO:58 is the nucleotide sequence of vector KS427.

SEQ ID NO:59 is the nucleotide sequence of vector KS130.

SEQ ID NO:60 is the nucleotide sequence of vector KS432.

SEQ ID NO:61 is the nucleotide sequence of vector ARALO80,

SEQ ID NO:62 is the nucleotide sequence of primer D6 fwd.

SEQ ID NO:63 is the nucleotide sequence of primer D6 rev,

SEQ ID NO:64 is the nucleotide sequence of vector KS119.

SEQ ID NO:65 is the nucleotide sequence of vector KS430.

SEQ ID NO:66 is the nucleotide sequence of vector ARALO78.

SEQ ID NO:67 is the nucleotide sequence of vector KS428.

SEQ ID NO:68 is the nucleotide sequence of vector KS429.

SEQ ID NO:69 is the nucleotide sequence of vector ARALO77.

SEQ ID NO:70 is the nucleotide sequence of vector KS431.

SEQ ID NO:71 is the nucleotide sequence of vector ARALO79.

SEQ ID NO:72 is the nucleotide sequence of the sucrose synthase 2-1 (BnSUS2-1) promoter from *Brassica napus* that is present in BN SUS2 prom1/PCR blunt.

SEQ ID NO:73 is the nucleotide sequence of the sucrose synthase 2-2 (BnSUS2-2) promoter from *Brassica napus* that is present in BN SUS2 prom2/PCR blunt.

The sequence descriptions and Sequence Listing attached hereto comply with the rules governing nucleotide and/or amino acid sequence disclosures in patent applications as set forth in 37 C.F.R. §1.821-1.825.

The Sequence Listing contains the one letter code for nucleotide sequence characters and the three letter codes for amino acids as defined in conformity with the IUPAC-IUBMB standards described in *Nucleic Acids Res.* 13:3021-3030 (1985) and in the *Biochemical J.* 219 (No. 2):345-373 (1984) which are herein incorporated by reference. The sym-

bols and format used for nucleotide and amino acid sequence data comply with the rules set forth in 37 C.F.R. §1.822.

#### DETAILED DESCRIPTION OF THE INVENTION

All patents, patent applications, and publications cited herein are incorporated by reference in their entirety.

As used herein and in the appended claims, the singular forms “a”, “an”, and “the” include plural reference unless the context clearly dictates otherwise. Thus, for example, reference to “a plant” includes a plurality of such plants, reference to “a cell” includes one or more cells and equivalents thereof known to those skilled in the art, and so forth.

Units, prefixes, and symbols may be denoted in their SI accepted form. Unless otherwise indicated, nucleic acids are written left to right in 5' to 3' orientation; amino acid sequences are written left to right in amino to carboxyl orientation, respectively. Numeric ranges recited within the specification are inclusive of the numbers defining the range and include each integer within the defined range. Amino acids may be referred to herein by either commonly known three letter symbols or by the one-letter symbols recommended by the IUPAC-IUB Biochemical Nomenclature Commission. Nucleotides, likewise, may be referred to by their commonly accepted single-letter codes. Unless otherwise provided for, software, electrical, and electronics terms as used herein are as defined in The New IEEE Standard Dictionary of Electrical and Electronics Terms (5<sup>th</sup> edition, 1993). The terms defined below are more fully defined by reference to the specification as a whole.

In the context of this disclosure, a number of terms and abbreviations are used. The following definitions are provided.

The term “ODP1” refers to an ovule development protein 1 that is involved with increasing oil content.

The term “sucrose synthase” (SUS) refers to an enzyme used in carbohydrate metabolism that catalyzes the reversible conversion of sucrose and uridine diphosphate (UDP) to UDP-glucose and fructose in vitro. The terms “*Arabidopsis* sucrose synthase 2”, “AtSuSy” and “AtSUS2”) are used interchangeably herein. The *Arabidopsis* sucrose synthase 2 gene is from genomic locus At5g49190.

The term “germination” refers to the initial stages in the growth of a seed to form a seedling.

The term “recombinant” refers to an artificial combination of two otherwise separated segments of sequence, e.g., by chemical synthesis or by the manipulation of isolated segments of nucleic acids by genetic engineering techniques.

The terms “recombinant construct”, “expression construct”, “chimeric construct”, “construct”, and “recombinant DNA construct” are used interchangeably herein. A recombinant construct comprises an artificial combination of nucleic acid fragments, e.g., regulatory and coding sequences that are not found together in nature. For example, a chimeric construct may comprise regulatory sequences and coding sequences that are derived from different sources, or regulatory sequences and coding sequences derived from the same source, but arranged in a manner different than that found in nature. Such a construct may be used by itself or may be used in conjunction with a vector. If a vector is used, then the choice of vector is dependent upon the method that will be used to transform host cells as is well known to those skilled in the art. For example, a plasmid vector can be used. The skilled artisan is well aware of the genetic elements that must be present on the vector in order to successfully transform, select and propagate host cells comprising any of the isolated nucleic acid fragments of the invention. The skilled artisan

will also recognize that different independent transformation events will result in different levels and patterns of expression (Jones et al., *EMBO J.* 4:2411-2418 (1985); De Almeida et al., *Mol. Gen. Genetics* 218:78-86 (1989)), and thus that multiple events must be screened in order to obtain lines displaying the desired expression level and pattern. Such screening may be accomplished by Southern analysis of DNA, Northern analysis of mRNA expression, immunoblotting analysis of protein expression, or phenotypic analysis, among others.

This construct may comprise any combination of deoxyribonucleotides, ribonucleotides, and/or modified nucleotides. The construct may be transcribed to form an RNA, wherein the RNA may be capable of forming a double-stranded RNA and/or hairpin structure. This construct may be expressed in the cell, or isolated or synthetically produced. The construct may further comprise a promoter, or other sequences which facilitate manipulation or expression of the construct.

As used herein, “encodes” or “encoding” refers to a DNA sequence which can be processed to generate an RNA and/or polypeptide.

As used herein, “expression” or “expressing” refers to production of a functional product, such as, the generation of an RNA transcript from an introduced construct, an endogenous DNA sequence, or a stably incorporated heterologous DNA sequence. The term may also refer to a polypeptide produced from an mRNA generated from any of the above DNA precursors. Thus, expression of a nucleic acid fragment may refer to transcription of the nucleic acid fragment (e.g., transcription resulting in mRNA or other functional RNA) and/or translation of RNA into a precursor or mature protein (polypeptide).

As used herein, “heterologous” with respect to a sequence means a sequence that originates from a foreign species, or, if from the same species, is substantially modified from its native form in composition and/or genomic locus by deliberate human intervention. For example, with respect to a nucleic acid, it can be a nucleic acid that originates from a foreign species, or is synthetically designed, or, if from the same species, is substantially modified from its native form in composition and/or genomic locus by deliberate human intervention. A heterologous protein may originate from a foreign species or, if from the same species, is substantially modified from its original form by deliberate human intervention.

“Plant” includes reference to whole plants, plant organs, plant tissues, seeds and plant cells and progeny of same. Plant cells include, without limitation, cells from seeds, suspension cultures, embryos, meristematic regions, callus tissue, leaves, roots, shoots, gametophytes, sporophytes, pollen, and microspores.

The term “plant parts” includes differentiated and undifferentiated tissues including, but not limited to the following: roots, stems, shoots, leaves, pollen, seeds, tumor tissue and various forms of cells and culture (e.g., single cells, protoplasts, embryos and callus tissue). The plant tissue may be in plant or in a plant organ, tissue or cell culture.

The term “plant organ” refers to plant tissue or group of tissues that constitute a morphologically and functionally distinct part of a plant.

“Progeny” comprises any subsequent generation of a plant. Progeny will inherit, and stably segregate, genes and transgenes from its parent plant(s).

The term “introduced” means providing a nucleic acid (e.g., expression construct) or protein into a cell. Introduced includes reference to the incorporation of a nucleic acid into a eukaryotic or prokaryotic cell where the nucleic acid may be incorporated into the genome of the cell, and includes reference to the transient provision of a nucleic acid or protein to

the cell. Introduced includes reference to stable or transient transformation methods, as well as sexually crossing. Thus, "introduced" in the context of inserting a nucleic acid fragment (e.g., a recombinant DNA construct/expression construct) into a cell, means "transfection" or "transformation" or "transduction" and includes reference to the incorporation of a nucleic acid fragment into a eukaryotic or prokaryotic cell where the nucleic acid fragment may be incorporated into the genome of the cell (e.g., chromosome, plasmid, plastid or mitochondrial DNA), converted into an autonomous replicon, or transiently expressed (e.g., transfected mRNA).

The term "genome" as it applies to a plant cells encompasses not only chromosomal DNA found within the nucleus, but organelle DNA found within subcellular components (e.g., mitochondrial, plastid) of the cell.

The term "isolated" refers to material, such as a nucleic acid or a protein, which is: (1) substantially or essentially free from components which normally accompany or interact with the material as found in its naturally occurring environment or (2) if the material is in its natural environment, the material has been altered by deliberate human intervention to a composition and/or placed at a locus in the cell other than the locus native to the material.

As used herein, "nucleic acid" means a polynucleotide and includes single or double-stranded polymer of deoxyribonucleotide or ribonucleotide bases. Nucleic acids may also include fragments and modified nucleotides. Thus, the terms "polynucleotide", "nucleic acid sequence", "nucleotide sequence" or "nucleic acid fragment" are used interchangeably and is a polymer of RNA or DNA that is single- or double-stranded, optionally containing synthetic, non-natural or altered nucleotide bases. Nucleotides (usually found in their 5'-monophosphate form) are referred to by their single letter designation as follows: "A" for adenylate or deoxyadenylate (for RNA or DNA, respectively), "C" for cytidylate or deoxycytidylate, "G" for guanylate or deoxyguanylate, "U" for uridylate, "T" for deoxythymidylate, "R" for purines (A or G), "Y" for pyrimidines (C or T), "K" for G or T, "H" for A or C or T, "I" for inosine, and "N" for any nucleotide.

The terms "subfragment that is functionally equivalent" and "functionally equivalent subfragment" are used interchangeably herein. These terms refer to a portion or subsequence of an isolated nucleic acid fragment in which the ability to alter gene expression or produce a certain phenotype is retained whether or not the fragment or subfragment encodes an active enzyme. For example, the fragment or subfragment can be used in the design of chimeric genes to produce the desired phenotype in a transformed plant. Chimeric genes can be designed for use in suppression by linking a nucleic acid fragment or subfragment thereof, whether or not it encodes an active enzyme, in the sense or antisense orientation relative to a plant promoter sequence.

The term "conserved domain" or "motif" means a set of amino acids conserved at specific positions along an aligned sequence of evolutionarily related proteins. While amino acids at other positions can vary between homologous proteins, amino acids that are highly conserved at specific positions indicate amino acids that are essential in the structure, the stability, or the activity of a protein. Because they are identified by their high degree of conservation in aligned sequences of a family of protein homologues, they can be used as identifiers, or "signatures", to determine if a protein with a newly determined sequence belongs to a previously identified protein family.

The terms "homology", "homologous", "substantially similar" and "corresponding substantially" are used interchangeably herein. They refer to nucleic acid fragments

wherein changes in one or more nucleotide bases do not affect the ability of the nucleic acid fragment to mediate gene expression or produce a certain phenotype. These terms also refer to modifications of the nucleic acid fragments of the instant invention such as deletion or insertion of one or more nucleotides that do not substantially alter the functional properties of the resulting nucleic acid fragment relative to the initial, unmodified fragment. It is therefore understood, as those skilled in the art will appreciate, that the invention encompasses more than the specific exemplary sequences.

"Sequence identity" or "identity" in the context of nucleic acid or polypeptide sequences refers to the nucleic acid bases or amino acid residues in two sequences that are the same when aligned for maximum correspondence over a specified comparison window.

Thus, "percentage of sequence identity" refers to the value determined by comparing two optimally aligned sequences over a comparison window, wherein the portion of the polynucleotide or polypeptide sequence in the comparison window may comprise additions or deletions (i.e., gaps) as compared to the reference sequence (which does not comprise additions or deletions) for optimal alignment of the two sequences. The percentage is calculated by determining the number of positions at which the identical nucleic acid base or amino acid residue occurs in both sequences to yield the number of matched positions, dividing the number of matched positions by the total number of positions in the window of comparison and multiplying the results by 100 to yield the percentage of sequence identity. Useful examples of percent sequence identities include, but are not limited to, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, or 95%, or any integer percentage from 50% to 100%. These identities can be determined using any of the programs described herein.

Sequence alignments and percent identity or similarity calculations may be determined using a variety of comparison methods designed to detect homologous sequences including, but not limited to, the MegAlign™ program of the LASERGENE bioinformatics computing suite (DNASTAR Inc., Madison, Wis.). Within the context of this application it will be understood that where sequence analysis software is used for analysis, that the results of the analysis will be based on the "default values" of the program referenced, unless otherwise specified. As used herein "default values" will mean any set of values or parameters that originally load with the software when first initialized.

The "Clustal V method of alignment" corresponds to the alignment method labeled Clustal V (described by Higgins and Sharp, *CABIOS*. 5:151-153 (1989); Higgins, D. G. et al. (1992) *Comput. Appl. Biosci.* 8:189-191) and found in the MEGALIGN™ program of the LASERGENE bioinformatics computing suite (DNASTAR Inc., Madison, Wis.). For multiple alignments, the default values correspond to GAP PENALTY=10 and GAP LENGTH PENALTY=10. Default parameters for pairwise alignments and calculation of percent identity of protein sequences using the Clustal method are KTUPLE=1, GAP PENALTY=3, WINDOW=5 and DIAGONALS SAVED=5. For nucleic acids these parameters are KTUPLE=2, GAP PENALTY=5, WINDOW=4 and DIAGONALS SAVED=4. After alignment of the sequences using the Clustal V program, it is possible to obtain a "percent identity" by viewing the "sequence distances" table in the same program.

"BLASTN method of alignment" is an algorithm provided by the National Center for Biotechnology Information (NCBI) to compare nucleotide sequences using default parameters.

It is well understood by one skilled in the art that many levels of sequence identity are useful in identifying polypeptides, from other species, wherein such polypeptides have the same or similar function or activity. Useful examples of percent identities include, but are not limited to, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, or 95%, or any integer percentage from 50% to 100%. Indeed, any integer amino acid identity from 50% to 100% may be useful in describing the present invention, such as 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99%. Also, of interest is any full-length or partial complement of this isolated nucleotide fragment.

"Gene" refers to a nucleic acid fragment that expresses a specific protein, including regulatory sequences preceding (5' non-coding sequences) and following (3' non-coding sequences) the coding sequence. "Native gene" refers to a gene as found in nature with its own regulatory sequences. "Chimeric gene" refers to any gene that is not a native gene, comprising regulatory and coding sequences that are not found together in nature. Accordingly, a chimeric gene may comprise regulatory sequences and coding sequences that are derived from different sources, or regulatory sequences and coding sequences derived from the same source, but arranged in a manner different than that found in nature. A "foreign" gene refers to a gene not normally found in the host organism, but that is introduced into the host organism by gene transfer. Foreign genes can comprise native genes inserted into a non-native organism, or chimeric genes. A "transgene" is a gene that has been introduced into the genome by a transformation procedure.

The term "genome" as it applies to a plant cells encompasses not only chromosomal DNA found within the nucleus, but organelle DNA found within subcellular components (e.g., mitochondrial, plastid) of the cell.

"Coding sequence" refers to a DNA sequence that codes for a specific amino acid sequence. "Regulatory sequences" refer to nucleotide sequences located upstream (5' non-coding sequences), within, or downstream (3' non-coding sequences) of a coding sequence, and which influence the transcription, RNA processing or stability, or translation of the associated coding sequence. Regulatory sequences may include, but are not limited to: promoters, translation leader sequences, introns, polyadenylation recognition sequences, RNA processing sites, effector binding sites and stem-loop structures.

"Promoter" refers to a DNA sequence capable of controlling the expression of a coding sequence or functional RNA. The promoter sequence consists of proximal and more distal upstream elements, the latter elements often referred to as enhancers. Accordingly, an "enhancer" is a DNA sequence that can stimulate promoter activity, and may be an innate element of the promoter or a heterologous element inserted to enhance the level or tissue-specificity of a promoter. Promoters may be derived in their entirety from a native gene, or be composed of different elements derived from different promoters found in nature, or even comprise synthetic DNA segments. It is understood by those skilled in the art that different promoters may direct the expression of a gene in different tissues or cell types, or at different stages of development, or in response to different environmental conditions. It is further recognized that since in most cases the exact boundaries of regulatory sequences have not been completely defined, DNA fragments of some variation may have identical promoter activity. Promoters that cause a gene to be

expressed in most cell types at most times are commonly referred to as "constitutive promoters". New promoters of various types useful in plant cells are constantly being discovered; numerous examples may be found in the compilation by Okamuro, J. K., and Goldberg, R. B. *Biochemistry of Plants* 15:1-82 (1989).

"Functional variants" of the regulatory sequences (e.g., promoters) are also encompassed by the compositions of the present invention. Functional variants include, for example, the native regulatory sequences of the invention having one or more nucleotide substitutions, deletions or insertions. Functional variants of the invention may be created by site-directed mutagenesis, induced mutation, or may occur as allelic variants (polymorphisms).

As used herein, a "functional fragment" of a regulatory sequence (e.g. a promoter) is a functional variant formed by one or more deletions from a larger regulatory element. For example, the 5' portion of a sequence with promoter activity may be deleted without abolishing promoter activity, as described by Zhu et al., *Plant Cell* 7:1681-1689 (1995). Such variants should retain promoter activity, particularly the ability to drive expression in seed or seed tissues. Activity can be measured by Northern blot analysis, reporter activity measurements when using transcriptional fusions, and the like. See, for example, Sambrook et al. (1989) *Molecular Cloning: A Laboratory Manual* (2<sup>nd</sup> ed., Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.).

"Translation leader sequence" refers to a polynucleotide sequence located between the promoter sequence of a gene and the coding sequence. The translation leader sequence is present in the fully processed mRNA upstream of the translation start sequence. The translation leader sequence may affect processing of the primary transcript to mRNA, mRNA stability or translation efficiency. Examples of translation leader sequences have been described (Turner, R. and Foster, G. D., *Mol. Biotechnol.* 3:225-236 (1995)).

"3' non-coding sequences", "transcription terminator" or "termination sequences" refer to DNA sequences located downstream of a coding sequence and include polyadenylation recognition sequences and other sequences encoding regulatory signals capable of affecting mRNA processing or gene expression. The polyadenylation signal is usually characterized by affecting the addition of polyadenylic acid tracts to the 3' end of the mRNA precursor. The use of different 3' non-coding sequences is exemplified by Ingelbrecht, I. L., et al. *Plant Cell* 1:671-680 (1989).

"RNA transcript" refers to the product resulting from RNA polymerase-catalyzed transcription of a DNA sequence. When the RNA transcript is a perfect complementary copy of the DNA sequence, it is referred to as the primary transcript. A RNA transcript is referred to as the mature RNA when it is a RNA sequence derived from post-transcriptional processing of the primary transcript. "Messenger RNA" or "mRNA" refers to the RNA that is without introns and that can be translated into protein by the cell. "cDNA" refers to a DNA that is complementary to, and synthesized from, a mRNA template using the enzyme reverse transcriptase. The cDNA can be single-stranded or converted into double-stranded form using the Klenow fragment of DNA polymerase I. "Sense" RNA refers to RNA transcript that includes the mRNA and can be translated into protein within a cell or in vitro. "Antisense RNA" refers to an RNA transcript that is complementary to all or part of a target primary transcript or mRNA, and that blocks the expression of a target gene (U.S. Pat. No. 5,107,065). The complementarity of an antisense RNA may be with any part of the specific gene transcript, i.e., at the 5' non-coding sequence, 3' non-coding sequence,

introns, or the coding sequence. "Functional RNA" refers to antisense RNA, ribozyme RNA, or other RNA that may not be translated but yet has an effect on cellular processes. The terms "complement" and "reverse complement" are used interchangeably herein with respect to mRNA transcripts, and are meant to define the antisense RNA of the message.

The term "operably linked" refers to the association of nucleic acid sequences on a single nucleic acid fragment so that the function of one is regulated by the other. For example, a promoter is operably linked with a coding sequence when it is capable of regulating the expression of that coding sequence (i.e., the coding sequence is under the transcriptional control of the promoter). Coding sequences can be operably linked to regulatory sequences in a sense or antisense orientation. In another example, the complementary RNA regions of the invention can be operably linked, either directly or indirectly, 5' to the target mRNA, or 3' to the target mRNA, or within the target mRNA, or a first complementary region is 5' and its complement is 3' to the target mRNA.

Standard recombinant DNA and molecular cloning techniques used herein are well known in the art and are described more fully in Sambrook, J., Fritsch, E. F. and Maniatis, T. *Molecular Cloning: A Laboratory Manual*; Cold Spring Harbor Laboratory: Cold Spring Harbor, N.Y. (1989). Transformation methods are well known to those skilled in the art and are described infra.

"PCR" or "polymerase chain reaction" is a technique for the synthesis of large quantities of specific DNA segments and consists of a series of repetitive cycles (Perkin Elmer Cetus Instruments, Norwalk, Conn.). Typically, the double-stranded DNA is heat denatured, the two primers complementary to the 3' boundaries of the target segment are annealed at low temperature and then extended at an intermediate temperature. One set of these three consecutive steps is referred to as a "cycle".

The terms "plasmid", "vector" and "cassette" refer to an extra chromosomal element often carrying genes that are not part of the central metabolism of the cell, and usually in the form of circular double-stranded DNA fragments. Such elements may be autonomously replicating sequences, genome integrating sequences, phage or nucleotide sequences, linear or circular, of a single- or double-stranded DNA or RNA, derived from any source, in which a number of nucleotide sequences have been joined or recombined into a unique construction which is capable of introducing a promoter fragment and DNA sequence for a selected gene product along with appropriate 3' untranslated sequence into a cell. "Transformation cassette" refers to a specific vector containing a foreign gene and having elements in addition to the foreign gene that facilitates transformation of a particular host cell. "Expression cassette" refers to a specific vector containing a foreign gene and having elements in addition to the foreign gene that allow for enhanced expression of that gene in a foreign host (i.e., to a discrete nucleic acid fragment into which a nucleic acid sequence or fragment can be moved.)

The term "expression", as used herein, refers to the production of a functional end-product (e.g., a mRNA or a protein [either precursor or mature]).

"Stable transformation" refers to the transfer of a nucleic acid fragment into a genome of a host organism, including both nuclear and organellar genomes, resulting in genetically stable inheritance. In contrast, "transient transformation" refers to the transfer of a nucleic acid fragment into the nucleus, or DNA-containing organelle, of a host organism resulting in gene expression without integration or stable

inheritance. Host organisms containing the transformed nucleic acid fragments are referred to as "transgenic" organisms.

As used herein, "transgenic" refers to a plant or a cell which comprises within its genome a heterologous polynucleotide. Preferably, the heterologous polynucleotide is stably integrated within the genome such that the polynucleotide is passed on to successive generations. The heterologous polynucleotide may be integrated into the genome alone or as part of an expression construct. Transgenic is used herein to include any cell, cell line, callus, tissue, plant part or plant, the genotype of which has been altered by the presence of heterologous nucleic acid including those transgenics initially so altered as well as those created by sexual crosses or asexual propagation from the initial transgenic. The term "transgenic" as used herein does not encompass the alteration of the genome (chromosomal or extra-chromosomal) by conventional plant breeding methods or by naturally occurring events such as random cross-fertilization, non-recombinant viral infection, non-recombinant bacterial transformation, non-recombinant transposition, or spontaneous mutation.

The present invention concerns a recombinant DNA construct comprising a polynucleotide encoding an ODP1 polypeptide operably linked to a sucrose synthase 2 promoter wherein said construct increases oil content in the seeds of a cruciferous oilseed plant while maintaining normal germination and further wherein the amino acid sequence of said ODP1 polypeptide has at least 80% sequence identity to a sequence selected from the group consisting of SEQ ID NO:37, SEQ ID NO:39, and SEQ ID NO:41.

In another embodiment, the sequence identity can be at least 90% or 95%.

In another embodiment the ODP1 polypeptide comprises a sequence selected from the group consisting of SEQ ID NO:37, SEQ ID NO:39, and SEQ ID NO:41.

In another embodiment, the sucrose synthase 2 promoter comprises: (a) the nucleotide sequence of SEQ ID NO:43, SEQ ID NO:72 or SEQ ID NO:73; or (b) a nucleotide sequence comprising a functional fragment of the nucleotide sequence of SEQ ID NO:43, SEQ ID NO:72 or SEQ ID NO:73.

ODP1 is a member of the APETALA2 (AP2) family of proteins that play a role in a variety of biological events including, but not limited to, oil content. The AP2/ERF family of proteins is a plant-specific class of putative transcription factors that have been shown to regulate a wide-variety of developmental processes and are characterized by the presence of an AP2/ERF DNA binding domain. The AP2/ERF proteins have been subdivided into two distinct subfamilies based on whether they contain one (ERF subfamily) or two (AP2 subfamily) DNA binding domains.

Specifically, AP2 (APETALA2) and EREBPs (ethylene-responsive element binding proteins) are the prototypic members of a family of transcription factors unique to plants, whose distinguishing characteristic is that they contain the so-called AP2 DNA-binding domain. AP2/EREBP genes form a large multigene family, and they play a variety of roles throughout the plant life cycle. AP2/EREBP genes are key regulators of several developmental processes, including floral organ identity determination and leaf epidermal cell identity. In *Arabidopsis thaliana*, the homeotic gene APETALA2 (AP2) has been shown to control three salient processes during development: (1) the specification of flower organ identity throughout floral organogenesis (Jofuku et al., *Plant Cell* 6:1211-1225, 1994); (2) establishment of flower meristem identity (Irish and Sussex, *Plant Cell* 2:8:741-753, 1990); and (3) the temporal and spatial regulation of flower homeotic

gene activity (Drews et al., Cell 65:6:991-1002, 1991). DNA sequence analysis suggests that AP2 encodes a theoretical polypeptide of 432 aa, with a distinct 68 aa repeated motif termed the AP2 domain. This domain has been shown to be essential for AP2 functions and contains within the 68 aa, an eighteen amino acid core region that is predicted to form an amphipathic  $\alpha$ -helix (Jofuku et al., Plant Cell 6:1211-1225, 1994). Apt-like domain-containing transcription factors have been also been identified in both *Arabidopsis thaliana* (Okamura et al., Proc. Natl. Acad. Sci. USA 94:7076-7081, 1997) and in tobacco with the identification of the ethylene responsive element binding proteins (EREBPs) (Ohme-Takagi and Shinshi, Plant Cell 7:2:173-182, 1995). In *Arabidopsis*, these RAP2 (related to AP2) genes encode two distinct subfamilies of AP2 domain-containing proteins designated AP2-like and EREBP-like (Okamura et al., Proc. Natl. Acad. Sci. USA 94:7076-7081, 1997). In vitro DNA binding has not been shown to date using the RAP2 proteins. Based upon the presence of two highly conserved motifs YRG and RAYD within the AP2 domain, it has been proposed that binding DNA binding occurs in a manner similar to that of AP2 proteins.

In another embodiment, the present invention concerns a transgenic cruciferous oilseed plant comprising in its genome the recombinant DNA construct of the invention. Also of interest is a transgenic seed obtained from a transgenic plant as described herein, wherein said seed comprises in its genome a recombinant DNA construct of the invention.

In still another aspect, the present invention concerns a method for producing a transgenic cruciferous oilseed plant comprising transforming a cruciferous oilseed plant cell with a recombinant construct of the invention and regenerating a transgenic plant from the transformed plant cell.

This invention concerns a transgenic seed obtained from a transgenic plant made by a method of the invention, wherein said seed comprises in its genome a recombinant DNA construct of the invention.

In another aspect, the present invention concerns a method for increasing oil content in seeds of a transgenic cruciferous oilseed plant while maintaining normal germination, said method comprising:

- (a) transforming a cruciferous oilseed plant cell with a recombinant DNA construct comprising a polynucleotide encoding an ODP1 polypeptide, wherein the amino acid sequence of said ODP1 polypeptide has at least 80%, at least 90% or at least 95% sequence identity with a sequence selected from the group consisting of SEQ ID NO:37, SEQ ID NO:39, and SEQ ID NO:41, said sequence being operably linked to a seed specific promoter;
- (b) regenerating a transgenic cruciferous oilseed plant from the transformed cell of step (a), wherein said plant comprises the recombinant DNA construct;
- (c) obtaining a transgenic progeny plant derived from the transgenic cruciferous oilseed plant of step (b), wherein the transgenic progeny plant comprises in its genome the recombinant DNA construct;
- (d) assaying the transgenic progeny plant obtained from step (c) for oil level and germination; and
- (e) selecting those transgenic progeny plants having seeds with an increased level of oil and normal germination when compared to seeds obtained from a control cruciferous oilseed plant, wherein said control plant does not comprise the recombinant DNA construct.

Preferably, the ODP1 polypeptide is a maize ODP1 polypeptide and, more preferably, the amino acid sequence of the ODP1 polypeptide comprises the sequence of SEQ ID NO:37.

With respect to the seed specific promoter, it can be a sucrose synthase 2 promoter and preferably, the nucleotide sequence of sucrose synthase 2 promoter comprises: (a) the nucleotide sequence of SEQ ID NO:43, SEQ ID NO:72 or SEQ ID NO:73; or (b) a nucleotide sequence comprising a functional fragment of the nucleotide sequence of SEQ ID NO:43, SEQ ID NO:72 or SEQ ID NO:73.

The transgenic cruciferous oil seeds described herein of the invention can be processed to yield oil and/or seed by-products.

In another embodiment, the present invention concerns a recombinant DNA construct comprising a polynucleotide encoding a heterologous polypeptide operably linked to a sucrose synthase 2 promoter, wherein the sucrose synthase 2 promoter comprises: (a) the nucleotide sequence of SEQ ID NO:43, SEQ ID NO:72 or SEQ ID NO:73; (b) a nucleotide sequence comprising a functional fragment of the nucleotide sequence of SEQ ID NO:43, SEQ ID NO:72 or SEQ ID NO:73; or (c) a nucleotide sequence with at least 80%, at least 90% or at least 95% sequence identity to the nucleotide sequence of SEQ ID NO:43, SEQ ID NO:72 or SEQ ID NO:73; wherein the nucleotide sequence of (a), (b) or (c) has seed-specific promoter activity in a plant. The invention also concerns a transgenic plant, plant cell and seed comprising the recombinant DNA construct. The transgenic plant may be a transgenic cruciferous plant.

The nucleotide and deduced amino acid sequence of the canola SUS2 homolog transcript model are set forth as SEQ ID NO:44 and SEQ ID NO:45, respectively.

NCBI GI NO. 150912532 is the nucleotide sequence of the 5'-end of a *Brassica oleracea* cDNA.

SEQ ID NO:72 is the nucleotide sequence of the sucrose synthase 2-1 (BnSUS2-1) promoter from *Brassica napus* that is present in BN SUS2 prom1/PCR blunt. Comparison of SEQ ID NO:72 with SEQ ID NO:44 and NCBI GI NO. 150912532 indicate that nucleotide 427 is at or near the beginning of the 5'-Untranslated region of the canola SUS2 gene. Consequently, a fragment comprising nucleotides 1-426 of SEQ ID NO:72 would be expected to have seed-specific promoter activity in a plant.

SEQ ID NO:73 is the nucleotide sequence of the sucrose synthase 2-2 (BnSUS2-2) promoter from *Brassica napus* that is present in BN SUS2 prom2/PCR blunt. Comparison of SEQ ID NO:73 with SEQ ID NO:44 and NCBI GI NO. 150912532 indicate that nucleotide 1766 is at or near the beginning of the 5'-Untranslated region of the canola SUS2 gene. Consequently, a fragment comprising nucleotides 1-1765 of SEQ ID NO:73 would be expected to have seed-specific promoter activity in a plant.

The cruciferous oilseed plant (or seed) of any of the compositions or methods of the present invention can be canola or *Arabidopsis* or other plant species including but not limited to the following: *Barbarea vulgaris*, *Brassica campestris*, *Brassica carinata*, *Brassica elongata*, *Brassica fruticulosa*, *Brassica hirta*, *Brassica juncea*, *Brassica napus*, *Brassica narinosa*, *Brassica nigra*, *Brassica oleracea*, *Brassica perviridis*, *Brassica rapa*, *Brassica rupestris*, *Brassica septiceps*, *Brassica tournefortii*, *Brassica verna*, *Camelina sativa*, *Crambe abyssinica*, *Lepidium campestre*, *Raphanus sativus*, *Sinapis alba*.

Methods of isolating seed oils are well known in the art: (Young et al., Processing of Fats and Oils, In *The Lipid Handbook*, Gunstone et al., eds., Chapter 5 pp 253-257;

Chapman & Hall: London (1994)). Seed by-products include but are not limited to the following: meal, lecithin, gums, free fatty acids, pigments, soap, stearine, tocopherols, sterols and volatiles.

The production of edible vegetable oils including canola oil involves two overall processes, mechanical pressing and extraction, and further processing to remove impurities. The techniques used are similar for most vegetable oils produced from the seeds of plants. The crushing and extraction processes utilized by the canola industry today produce very little change to the fatty acid profile of the oil and the nutritional qualities of the meal.

For example, canola seeds are crushed into two component parts, oil and meal, which are then further manufactured into a wide variety of products. Further manufacturing, called refining, improves the color, flavor and shelf life of canola oil.

Canola oil is extracted in several stages. The first stage in processing canola is to roll or flake the seed. This ruptures cells and makes the oil easier to extract. Next the flaked or rolled seeds are cooked and subjected to a mild pressing process which removes some of the oil and compresses the seeds into large chunks called "cake fragments." The cake fragments undergo further processing to remove most of the remaining oil. The oil extracted during each step is combined. The oil is then subjected to processing according to the end product requirements. Different treatments are used to process salad oils, margarines, and shortenings.

Specifically, canola seed is cleaned by a number of different methods including air aspiration, indent cylinder cleaning, sieve screening, or a combination of these. Cleaning ensures that the seed is free of extraneous plant and other foreign material which is referred to in the industry as "dockage". Seed generally contains less than 2.5% dockage following the cleaning process. Seed that has been cleaned is ready for subsequent crushing into canola oil and meal.

Seed which will be processed for oil and meal is preconditioned using mild heat treatment, and moisture is then adjusted to improve subsequent oil extraction. Following preconditioning, canola seed is next crushed and flaked and then heated slightly. These processes help to maximize oil recovery. The canola flakes are then "prepressed" in screw presses or expellers to reduce the oil content from about 42% in the seed (on an 8% moisture basis) to between 16-20%. Screw pressing also compresses the flakes into more dense cakes (called "press cake") which facilitates oil extraction.

Press cake which results from seed processing is next subjected to one of two types of oil extraction to remove much of the remaining oil. Oil may be extracted using either hexane ("solvent") extraction or by "cold-pressing" (also referred to as "expeller pressing"). The end-market into which the oil is sold generally dictates which form of extraction will be used. Hexane is the extraction medium used for the bulk of canola oil which is sold into the commodity grocery chain market as well as to the food industry. Cold-pressed canola oil represents a much smaller volume sold to consumers and is generally marketed in specialty food stores. Both extraction processes result in an oil essentially bland in taste, light yellow in color, and with excellent nutritional and stability properties.

Hexane extraction reduces the oil content of the press cake to very low levels. Oil recovery from canola seed is approximately 96% when this form of extraction is used. This is accomplished by maximizing contact of the hexane with the press cake through a series of soakings or washings. Residual hexane in the extracted press cake and oil is easily removed by evaporation at low temperature. Solvent residues in oils and

meals, when produced in accordance with good manufacturing practice, can be said to be truly insignificant.

The oil which is produced during the extraction process is referred to as "crude oil" because it contains various compounds which must be removed to ensure a product with good stability and shelf-life. These impurities include phospholipids, mucilaginous gums, free fatty acids, color pigments and fine meal particles. Different methods are used to remove these by-products including water precipitation or organic acids in combination with water. Once removed, these by-products are added to the canola meal fraction in order to increase its feeding value (energy) and make it an even more nutritious product.

Following water precipitation and/or organic acid processing, the oil will still contain color compounds which, if not removed would make it unattractive to the consumer and also reduce its stability. These compounds are extracted through a process called bleaching. In contrast to what may be implied by the term, bleaching does not involve the use of harsh chemicals. Instead, during the bleaching process, the oil is moved through a natural, diatomaceous clay to remove color compounds and other by-products.

Deodorization is the final step in the refining of all vegetable oils, including canola. Deodorization involves the use of steam distillation with the objective being the removal of any residual compounds which, if retained, could impart an adverse odor and taste to the oil. The oil produced is referred to as "refined oil".

In still another embodiment, this invention concerns a transgenic progeny plant obtained from the plant of claim 7 or 12, wherein said transgenic progeny plant comprises the recombinant DNA construct.

There are a variety of methods for the regeneration of plants from plant tissue. The particular method of regeneration will depend on the starting plant tissue and the particular plant species to be regenerated. The regeneration, development and cultivation of plants from single plant protoplast transformants or from various transformed explants is well known in the art (Weissbach and Weissbach, In: *Methods for Plant Molecular Biology*, (Eds.), Academic: San Diego, Calif. (1988)). This regeneration and growth process typically includes the steps of selection of transformed cells and culturing those individualized cells through the usual stages of embryonic development through the rooted plantlet stage. Transgenic embryos and seeds are similarly regenerated. The resulting transgenic rooted shoots are thereafter planted in an appropriate plant growth medium such as soil. Preferably, the regenerated plants are self-pollinated to provide homozygous transgenic plants. Otherwise, pollen obtained from the regenerated plants is crossed to seed-grown plants of agronomically important lines. Conversely, pollen from plants of these important lines is used to pollinate regenerated plants. A transgenic plant of the present invention containing a desired polypeptide is cultivated using methods well known to one skilled in the art.

Normal germination of transgenic plant seed is defined as germination frequency that is very similar to the germination frequency of seed of the untransformed variety under produced under identical conditions.

In addition to the above discussed procedures, practitioners are familiar with the standard resource materials which describe specific conditions and procedures for: the construction, manipulation and isolation of macromolecules (e.g., DNA molecules, plasmids, etc.); the generation of recombinant DNA fragments and recombinant expression constructs; and, the screening and isolating of clones. See, for example: Sambrook et al., *Molecular Cloning: A Laboratory Manual*,

## 21

Cold Spring Harbor: NY (1989); Maliga et al., *Methods in Plant Molecular Biology*, Cold Spring Harbor: NY (1995); Birren et al., *Genome Analysis: Detecting Genes*, Vol. 1, Cold Spring Harbor: NY (1998); Birren et al., *Genome Analysis: Analyzing DNA*, Vol. 2, Cold Spring Harbor: NY (1998); *Plant Molecular Biology: A Laboratory Manual*, eds. Clark, Springer: NY (1997).

Examples of cruciferous oilseed plants that can be used to practice the invention include, but are not limited to, *Brassica* species, and *Arabidopsis thaliana*.

Assays for gene expression based on the transient expression of cloned nucleic acid constructs have been developed by introducing the nucleic acid molecules into plant cells by polyethylene glycol treatment, electroporation, or particle bombardment (Marcotte et al., *Nature* 335:454-457 (1988); Marcotte et al., *Plant Cell* 1:523-532 (1989); McCarty et al., *Cell* 66:895-905 (1991); Hattori et al., *Genes Dev.* 6:609-618 (1992); Goff et al., *EMBO J.* 9:2517-2522 (1990)).

Transient expression systems may be used to functionally dissect gene constructs (see generally, Maliga et al., *Methods in Plant Molecular Biology*, Cold Spring Harbor Press (1995)). It is understood that any of the nucleic acid molecules of the present invention can be introduced into a plant cell in a permanent or transient manner in combination with other genetic elements such as vectors, promoters, enhancers etc.

In addition to the above discussed procedures, practitioners are familiar with the standard resource materials which describe specific conditions and procedures for the construction, manipulation and isolation of macromolecules (e.g., DNA molecules, plasmids, etc.), generation of recombinant organisms and the screening and isolating of clones, (see for example, Sambrook et al., *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Press (1989); Maliga et al., *Methods in Plant Molecular Biology*, Cold Spring Harbor Press (1995); Birren et al., *Genome Analysis: Detecting Genes*, 1, Cold Spring Harbor, N.Y. (1998); Birren et al., *Genome Analysis: Analyzing DNA*, 2, Cold Spring Harbor, N.Y. (1998); *Plant Molecular Biology: A Laboratory Manual*, eds. Clark, Springer, New York (1997)).

## EXAMPLES

The present invention is further defined in the following Examples, in which parts and percentages are by weight and degrees are Celsius, unless otherwise stated. It should be understood that these Examples, while indicating preferred embodiments of the invention, are given by way of illustration only. From the above discussion and these Examples, one skilled in the art can ascertain the essential characteristics of this invention, and without departing from the spirit and scope thereof, can make various changes and modifications of the invention to adapt it to various usages and conditions. Thus, various modifications of the invention in addition to those shown and described herein will be apparent to those skilled in the art from the foregoing description. Such modifications are also intended to fall within the scope of the appended claims.

The meaning of abbreviations is as follows: "seq" means second(s), "min" means minute(s), "h" means hour(s), "d" means day(s), "μL" means microliter(s), "mL" means milliliter(s), "L" means liter(s), "μM" means micromolar, "mM" means millimolar, "M" means molar, "mmol" means millimole(s), "μmole" mean micromole(s), "g" means gram(s),

## 22

"μg" means microgram(s), "ng" means nanogram(s), "U" means unit(s), "bp" means base pair(s) and "kB" means kilobase(s).

## Example 1

Construction of Vector pZBL120×KS336 for Expression of a *Zea mays* ODP1 Under Control of a Beta-Conglycinin Promoter

Plasmid pKS332 was constructed via a number of different intermediate vectors. The *AscI* cassette containing Kti3 Promoter::NotI::Kti3 Terminator from pKS121 (PCT Application No. WO 02/00904) was blunt-end cloned into the NotI (filled-in) site on pBLUESCRIPT® II SK+ (Stratagene) to give pKS121/BS (SEQ ID NO:1). The *NcoI*/NotI fragment from expression vector pDsRed-Express (Clontech) was blunt-end cloned into the NotI (filled-in) site of pKS121/BS to give pDsRedxKS121/BS (SEQ ID NO:2). The *Bam*HI cassette containing Kti3 Promoter::DsRed::Kti3 Terminator in pDS-REDxKS121/BS (SEQ ID NO:1) was ligated into the *Bam*HI site of pKS123 (PCT Application No. WO 02/08269) to give pKS332 (SEQ ID NO:3). A DNA fragment encoding the ODP1 polypeptide from maize, Zm-ODP1, described in U.S. Pat. No. 7,157,621, was synthesized by PCR with primers to introduce NotI sites at both ends. Applicants cDNA clone *cde1c.pk003.o22* (SEQ ID NO:319 in U.S. Pat. No. 7,157,621) was used as template in a PCR reaction using primers MWG345 (SEQ ID NO:4) and MWG346 (SEQ ID NO:5). The resulting PCR product was digested with NotI restriction enzyme and ligated into the NotI site of pKS332 to give pKS336 (SEQ ID NO:6). Plasmid pKS336 contains the ZM-ODP1 protein-coding region of cDNA clone *cde1c.pk003.o22* fused at its 5' terminus with the promoter of the soybean gene for the α'-subunit of β-conglycinin (Beachy et al. (1985) *EMBO J.* 4:3047-3053) and at its 3' end with the terminator sequence from the phaseolin gene of common bean, *Phaseolus vulgaris* (Doyle et al. (1986) *J. Biol. Chem.* 261:9228-9238). The β-conglycinin promoter directs strong seed-specific expression of transgenes in transformed plants.

A 5.9 kb DNA fragment containing the ZM-ODP1 and DsRed expression cassettes was excised from KS336 using the restriction enzyme *AscI* and the ends were filled-in with T4 DNA polymerase (Promega, Madison, USA). This fragment was ligated to linearized DNA of the *Agrobacterium tumefaciens* binary vector pZBL120, which had been linearized with *EcoRI* and *Bam*HI and the ends filled-in, to give pZBL120×KS336. The T-DNA of the plant transformation vector pZBL120×KS336 is set forth as SEQ ID NO:7.

It is noted that the binary vector pZBL120 is identical to the pZBL1 binary vector (American Type Culture Collection Accession No. 209128) described in U.S. Pat. No. 5,968,793, except the NOS promoter was replaced with a 963 bp 35S promoter (NCBI Accession No. V00141; also known as NCBI General Identifier No. 58821) from nucleotide 6494 to 7456 in the NOS Promoter::nptII::OCS Terminator cassette. The new 35S Promoter::nptII::OCS Terminator cassette serves as a kanamycin (Kan) resistance plant selection marker in pZBL120.

## Example 2

Generation and Analysis of Oil Content of Transgenic *Arabidopsis* Lines Containing a Beta-Conglycinin Promoter::ZM-ODP1::Phaseolin Terminator Expression Cassette

Plasmid DNA of pZBL120×KS336, containing the beta-conglycinin promoter::ZM-ODP1::phaseolin terminator

## 23

expression cassette, was introduced into *Agrobacterium tumefaciens* NTL4 (Luo et al, *Molecular Plant-Microbe Interactions* (2001) 14(1):98-103) by electroporation. Briefly, 1 µg plasmid DNA was mixed with 100 µL of electro-competent cells on ice. The cell suspension was transferred to a 100 µL electroporation cuvette (1 mm gap width) and electroporated using a BIORAD electroporator set to 1 kV, 4000 and 25 µF. Cells were transferred to 1 mL LB medium and incubated for 2 h at 30° C. Cells were plated onto LB medium containing 50 µg/mL kanamycin. Plates were incubated at 30° C. for 60 h. Recombinant *Agrobacterium* cultures (500 mL LB, 50 µg/mL kanamycin) were inoculated from single colonies of transformed *agrobacterium* cells and grown at 30° C. for 60 h. Cells were harvested by centrifugation (5000×g, 10 min) and resuspended in 1 L of 5% (W/V) sucrose containing 0.05% (V/V) Silwet. *Arabidopsis* plants were grown in soil at a density of 30 plants per 100 cm<sup>2</sup> pot in METRO-MIX® 360 soil mixture for 4 weeks (22° C., 16 h light/8 h dark, 100 µE m<sup>-2</sup>s<sup>-1</sup>). Plants were repeatedly dipped into the *Agrobacterium* suspension harboring the binary vector pZBL120×KS336 and kept in a dark, high humidity environment for 24 h. Plants were grown for three to four weeks under standard plant growth conditions described above and plant material was harvested and dried for one week at ambient temperatures in paper bags. Seeds were harvested using a 0.425 mm mesh brass sieve.

Cleaned *Arabidopsis* seeds (2 grams, corresponding to about 100,000 seeds) were sterilized by washes in 45 mL of 80% ethanol, 0.01% TRITON® X-100, followed by 45 mL of 30% (V/V) household bleach in water, 0.01% TRITON® X-100 and finally by repeated rinsing in sterile water. Aliquots of 20,000 seeds were transferred to square plates (20×20 cm) containing 150 mL of sterile plant growth medium comprised of 0.5×MS salts, 0.53% (W/V) sorbitol, 0.05 MES/KOH (pH 5.8), 200 µg/mL TIMENTIN®, and 50 µg/mL kanamycin solidified with 10 g/L agar. Homogeneous dispersion of the seed on the medium was facilitated by mixing the aqueous seed suspension with an equal volume of melted plant growth medium. Plates were incubated under standard growth conditions for ten days. Kanamycin-resistant seedlings were transferred to plant growth medium without selective agent and grown for one week before transfer to soil. Plants were grown to maturity and T2 seeds were harvested and plated on selective media containing kanamycin. Approximately 100 events were generated in this manner. Wild-type (WT) control plants were grown in the same flat containing pZBL120×KS336 T1 plants. T2 seed were harvested and oil content was measured by NMR as follows.

## NMR Based Analysis of Seed Oil Content:

Seed oil content was determined using a Maran Ultra NMR analyzer (Resonance Instruments Ltd, Whitney, Oxfordshire, UK). Samples (e.g., batches of *Arabidopsis* seed ranging in weight between 5 and 200 mg) were placed into pre-weighed 2 mL polypropylene tubes (Corning Inc, Corning N.Y., USA; Part no. 430917) previously labeled with unique bar code identifiers. Samples were then placed into 96 place carriers and processed through the following series of steps by an ADEPT COBRA 600™ SCARA robotic system:

1. pick up tube (the robotic arm was fitted with a vacuum pickup device);
2. read bar code;
3. expose tube to antistatic device (ensured that *Arabidopsis* seed were not adhering to the tube walls);
4. weigh tube (containing the sample), to 0.0001 g precision;

## 24

5. take NMR reading; measured as the intensity of the proton spin echo 1 msec after a 22.95 MHz signal had been applied to the sample (data was collected for 32 NMR scans per sample);
6. return tube to rack; and
7. repeat process with next tube.

Bar codes, tubes weights and NMR readings were recorded by a computer connected to the system. Sample weight was determined by subtracting the polypropylene tube weight from the weight of the tube containing the sample.

Seed oil content (on a % seed weight basis) of *Arabidopsis* seed was calculated as follows:

$$\text{mg oil} = (\text{NMR signal} - 2.1112) / 37.514;$$

$$\% \text{ oil} = [(\text{mg oil}) / 1000] / [\text{g of seed sample weight}] \times 100.$$

Prior to establishing this formula, *Arabidopsis* seed oil was extracted as follows. Approximately 5 g of mature *Arabidopsis* seed (cv Columbia) were ground to a fine powder using a mortar and pestle. The powder was placed into a 33×94 mm paper thimble (Ahlstrom #7100-3394; Ahlstrom, Mount Holly Springs, Pa., USA) and the oil extracted during approximately 40 extraction cycles with petroleum ether (BP 39.9-51.7° C.) in a Soxhlet apparatus. The extract was allowed to cool and the crude oil was recovered by removing the solvent under vacuum in a rotary evaporator. Calibration parameters were determined by precisely weighing 11 standard samples of partially purified *Arabidopsis* oil (samples contained 3.6, 6.3, 7.9, 9.6, 12.8, 16.3, 20.3, 28.2, 32.1, 39.9 and 60 mg of partially purified *Arabidopsis* oil) weighed to a precision of 0.0001 g into 2 mL polypropylene tubes (Corning Inc, Corning N.Y., USA; Part no. 430917) and subjecting them to NMR analysis. A calibration curve of oil content (% seed weight basis) to NMR value was established.

Seed oil content of most transgenic lines was increased when compared to oil content of seed collected from wild-type control plants grown in the same flat. The phenotype of two representative transgenic lines, C00536 and C00576, are described below in detail. Kanamycin-resistant T2 seedlings were transferred from selective growth media to soil. For C00536, thirteen T2 plants were grown with four wild-type (WT) control plants. For C00576 ten T2 plants were grown with seven WT control plants. Plants were grown to maturity, T3 seed were harvested from individual plants and subjected to oil quantitation by NMR.

Data are summarized in Table 1. Presence of the pZBL120×KS336 transgene is associated with an increase in oil content in transgenic T3 seed when compared *Arabidopsis* plants of identical genetic background that lack the transgene.

TABLE 1

Oil Content of T3 Seed of pZBL120×KS336 Transgenics			
Exp	Event ID	Plant #	% Oil
1	C00536	1	45.7
1	C00536	2	45.1
1	C00536	3	45.0
1	C00536	4	44.6
1	C00536	5	44.0
1	C00536	6	43.7
1	C00536	7	43.5
1	C00536	8	42.8
1	C00536	9	42.7
1	C00536	10	42.0
1	C00536	11	42.0
1	C00536	12	41.9
1	C00536	13	39.9
1	C00536	AVG	43.3

TABLE 1-continued

Oil Content of T3 Seed of pZBL120xKS336 Transgenics			
Exp	Event ID	Plant #	% Oil
1	WT	1	39.5
1	WT	2	37.5
1	WT	3	37.0
1	WT	4	34.7
1	WT	AVG	37.2
2	C00576	1	48.0
2	C00576	2	47.9
2	C00576	3	45.9
2	C00576	4	45.3
2	C00576	5	44.5
2	C00576	6	43.7
2	C00576	7	43.6
2	C00576	8	42.1
2	C00576	9	41.9
2	C00576	10	41.0
2	C00576	AVG	44.4
2	WT	1	42.2
2	WT	2	40.9
2	WT	3	40.4
2	WT	4	39.3
2	WT	5	38.7
2	WT	6	38.0
2	WT	7	37.8
2	WT	AVG	39.6

Transgenic T3 seed selections that no longer segregated for the DsRed marker gene were identified by visual inspection using a suitable light source. For both events non-segregating transgenic seed were planted in soil alongside untransformed WT plants.

T4 seed were harvested from individual T3 plants and WT controls. Oil content was measured by NMR (Table 2). Presence of the pZBL120xKS336 transgene is associated with an increase in oil content in transgenic T4 seed when compared to *Arabidopsis* plants of identical genetic background that lack the transgene.

TABLE 2

Oil Content of T4 Seed of pZBL120xKS336 Transgenics			
Exp	Event ID	Plant #	% Oil
1	C00536	1	46.5
1	C00536	2	46.5
1	C00536	3	46.4
1	C00536	4	46.3
1	C00536	5	46.3
1	C00536	6	46.2
1	C00536	7	46.2
1	C00536	8	46.2
1	C00536	9	46.2
1	C00536	10	46.1
1	C00536	11	46.0
1	C00536	12	45.8
1	C00536	13	45.2
1	C00536	14	45.1
1	C00536	15	45.1
1	C00536	16	44.5
1	C00536	17	43.5
1	C00536	18	43.4
1	C00536	AVG	45.6
1	WT	1	44.8
1	WT	2	44.6
1	WT	3	42.3
1	WT	4	42.1
1	WT	5	42.0
1	WT	AVG	43.2
2	C00536	1	45.7
2	C00536	2	45.6
2	C00536	3	45.6
2	C00536	4	45.4

TABLE 2-continued

Oil Content of T4 Seed of pZBL120xKS336 Transgenics			
Exp	Event ID	Plant #	% Oil
2	C00536	5	45.4
2	C00536	6	45.4
2	C00536	7	45.4
2	C00536	8	45.4
2	C00536	9	45.4
2	C00536	10	45.1
2	C00536	11	45.1
2	C00536	12	45.0
2	C00536	13	44.8
2	C00536	14	44.7
2	C00536	15	44.6
2	C00536	16	44.5
2	C00536	17	43.5
2	C00536	18	43.1
2	C00536	AVG	45.0
2	WT	1	43.8
2	WT	2	43.3
2	WT	3	42.3
2	WT	4	41.8
2	WT	5	41.5
2	WT	6	40.2
2	WT	AVG	42.1
3	C00576	1	45.3
3	C00576	2	44.8
3	C00576	3	44.7
3	C00576	4	44.7
3	C00576	5	44.4
3	C00576	6	44.2
3	C00576	7	44.2
3	C00576	8	44.2
3	C00576	9	44.2
3	C00576	10	44.0
3	C00576	11	43.8
3	C00576	12	43.3
3	C00576	13	43.1
3	C00576	14	43.0
3	C00576	15	41.8
3	C00576	16	41.1
3	C00576	AVG	43.8
3	WT	1	43.8
3	WT	2	42.9
3	WT	3	42.4
3	WT	4	41.9
3	WT	5	41.6
3	WT	6	40.3
3	WT	7	37.5
3	WT	8	41.1
3	WT	AVG	41.4
4	C00576	1	46.6
4	C00576	2	46.4
4	C00576	3	46.3
4	C00576	4	46.2
4	C00576	5	46.2
4	C00576	6	46.2
4	C00576	7	46.2
4	C00576	8	45.7
4	C00576	9	45.7
4	C00576	10	45.6
4	C00576	11	45.6
4	C00576	12	45.4
4	C00576	13	45.4
4	C00576	14	45.1
4	C00576	15	45.0
4	C00576	16	44.3
4	C00576	17	44.2
4	C00576	AVG	45.7
4	WT	1	44.7
4	WT	2	44.6
4	WT	3	44.4
4	WT	4	43.7
4	WT	5	43.5
4	WT	6	42.2
4	WT	AVG	43.9

27

A total of five flats were planted using WT seed and homozygous T4 seed of C00536 and C00576. Twenty-four transgenic T4 plants were grown alongside twelve WT plants. Plants were grown to maturity. From each flat WT and transgenic seed were bulk-harvested. Oil content of bulk seed samples was measured by NMR (Table 3). Presence of the pZBL120×KS336 transgene is associated with an increase in oil content in transgenic T5 seed when compared to *Arabidopsis* plants of identical genetic background that lack the transgene.

Seed oil content in a given plant is a highly variable trait that responds strongly to plant growth conditions (Li Y, Beisson F, Pollard M, Ohlrogge J (2006) Oil content of *Arabidopsis* seeds: The influence of seed anatomy, light and plant-to-plant variation, *Phytochemistry* 67:904-915). It is therefore important that an increase in oil content associated with a particular strategy is observed in multiple environments, over several generations and under conditions that allow for maximal oil accumulation by isogenic control lines. The increase in oil content associated with presence of the pZBL120×KS336 transgene was consistently observed over three generations and in different growth chambers. The average oil increase associated with two different pZBL120×KS336 transgenic events was at least 2% points and as high as 3.6% points (i.e., an oil increase of as high as 8.5% compared to untransformed WT seed). This oil increase was observed under growth conditions in which untransformed *Arabidopsis* seed produced the expected levels of oil, indicating that oil seed storage lipid accumulation was operating at maximum levels.

TABLE 3

Oil Content of T5 Seed of pZBL120×KS336 Transgenics				
Flat ID	Event ID	Oil (%)	Δ Oil (%) Points)	Δ □Oil (%)
A	C00576	45.1	1.7	3.9
	WT	43.5		
B	C00576	46.4	1.9	4.2
	WT	44.5		
C	C00576	44.8	2.3	5.5
	WT	42.5		
D	C00576	45.5	2.0	4.7
	WT	43.4		
E	C00576	44.6	2.0	4.7
	WT	42.6		
AVG	C00576		2.0	4.6
A	C00536	45.9	3.3	7.8
	WT	42.6		
B	C00536	45.8	3.4	8.1
	WT	42.4		
C	C00536	46.7	4.7	11.2
	WT	42.0		
D	C00536	44.7	3.9	9.6
	WT	40.8		
E	C00536	46.2	2.6	6.0
	WT	43.5		
AVG	C00536		3.6	8.5

## Example 3

Construction of Vector pZBL120×KS333 for Expression of a *Momordica charantia* ODP1 Under Control of a Beta-Conglycinin Promoter

An ODP1 protein-coding region from balsam pear (*Momordica charantia*) described in detail in U.S. Pat. No. 7,157,621 was synthesized by PCR with primers to introduce NotI sites at both ends of the gene. Applicants cDNA clone

28

fdsIn.pk015.115 was used a template in the PCR reaction using primers MWG339 (SEQ ID NO:8) and MWG340 (SEQ ID NO:9). The resulting PCR product was digested with NotI restriction enzyme and ligated into the NotI site of pKS332 to give pKS333 (SEQ ID NO:10).

A 6.1 kb DNA fragment containing the MC-ODP1 and DsRed expression cassettes was excised from KS333 using the restriction enzyme AscI, the ends were filled-in with T4 DNA polymerase (Promega, Madison, USA) and the fragment was blunt-end ligated to DNA of the *Agrobacterium tumefaciens* binary vector pZBL120, which had been linearized with EcoRI and BamHI and the ends filled-in. The resulting plant transformation vector was designated pZBL120×KS333, and the T-DNA of this vector is set forth as SEQ III NO:11.

## Example 4

Construction of Vector pZBL120×KS334 for Expression of a *Glycine max* ODP1 Under Control of a Beta-Conglycinin Promoter

An ODP1 protein-coding region from soybean described in detail in U.S. Pat. No. 7,157,621 was synthesized by PCR with primers to introduce NotI sites at both ends of the gene. Applicants cDNA clone se3.pk0003.f5 was used as template in the PCR reaction using primers MWG341 (SEQ ID NO:12) and MWG342 (SEQ ID NO:13). The resulting PCR product was digested with NotI restriction enzyme and ligated into the NotI site of pKS332 to give pKS334 (SEQ ID NO:14).

A 6.1 kb DNA fragment containing the GM-ODP1 and DsRed expression cassettes was excised from KS334 using the restriction enzyme AscI, the ends were filled-in with T4 DNA polymerase (Promega, Madison, USA) and the fragment was blunt-end ligated to DNA of the *Agrobacterium tumefaciens* binary vector pZBL120, which had been linearized with EcoRI and BamHI and the ends filled-in. The resulting plant transformation vector was designated pZBL120×KS334, and the T-DNA of this vector is set forth as SEQ ID NO:15.

## Example 5

Generation of *Arabidopsis* Lines Transformed with *Momordica charantia* ODP1 or *Glycine max* ODP1 and Analysis of Seed Oil Content

Binary vector constructs pZBL120×KS333 (*Momordica charantia* ODP1) and pZBL120×KS334 (*Glycine max* ODP1) were used for *Arabidopsis* transformation using the floral dip method as described above. Transgenic lines were selected on plant growth media containing kanamycin. 75 and 190 lines were generated with pZBL120×KS333 and pZBL120×KS334, respectively. T1 plants of all lines were grown with 13 untransformed WT plants in the same growth chamber. Plants were grown to maturity. Seed were harvested from individual plants and oil content was measured by NMR (TABLE 4)

TABLE 4

Oil Content of T2 seed of pZBL120xKS333 and pZBL120xKS334 Transgenics			
<i>Arabidopsis</i> Line	# of Plants	% Oil Range	Average % Oil
pZBL120xKS333	77	25.5-46.6	41.7
pZBL120xKS334	180	16.0-48.1	40.7
WT	13	31.9-43.2	39.1

T2 seed of two representative transgenic lines, 4445 (pZBL120xKS333) and 4485 (pZBL120xKS334), had an oil content of 45.1% and 45.2% respectively. T2 seed of these two lines were germinated on selective media, seedlings were transferred to soil, T2 plants were grown to maturity and T3 seed were harvested. After one more round of germination on selective media and seed production for each event five flats were planted with 24 kanamycin-resistant 4445 or 4485 seedlings and 12 WT seedlings. Plants were grown to maturity. From each flat WT and transgenic seed were bulk-harvested. Oil content of bulk seed samples was measured by NMR (Table 5). Presence of the pZBL120xKS333 or pZBL120xKS334 transgenes is associated with an increase in oil content in transgenic T5 seed when compared to *Arabidopsis* plants of identical genetic background that lack the transgene.

TABLE 5

Oil Content of T5 seed of pZBL120xKS333 and pZBL120xKS334 Transgenics					
Flat ID	Construct	Event ID	Oil (%)	$\Delta$ Oil (% Points)	$\Delta$ Oil (%)
A	pZBL120xKS333	4445	44.9	0.7	1.5
		WT	44.2		
B	pZBL120xKS333	4445	45.3	1.8	4.0
		WT	43.6		
C	pZBL120xKS333	4445	46.0	2.4	5.4
		WT	43.7		
D	pZBL120xKS333	4445	44.6	1.4	3.2
		WT	43.2		
E	pZBL120xKS333	4445	43.2	-0.6	-1.4
		WT	43.8		
AVG	pZBL120xKS333			1.1	2.5
A	pZBL120xKS334	4485	45.4	2.8	6.7
		WT	42.5		
B	pZBL120xKS334	4485	44.4	1.3	3.1
		WT	43.1		
C	pZBL120xKS334	4485	44.5	1.7	4.0
		WT	42.8		
D	pZBL120xKS334	4485	45.1	1.5	3.3
		WT	43.7		
E	pZBL120xKS334	4485	45.4	1.6	3.8
		WT	43.8		
AVG	pZBL120xKS334			1.8	4.2

The oil increase associated with presence of the *Momordica charantia* ODP1 transgene (pZBL120xKS333) is 1.1% points (i.e., an oil increase of 2.5% compared to untransformed WT seed).

The oil increase associated with presence of the *Glycine max* ODP1 transgene (pZBL120xKS334) is 1.8% points (i.e., an oil increase of 4.2% compared to untransformed WT seed).

#### Example 6

##### Compositional Analysis of *Arabidopsis* Seed Transformed with *Zea mays* ODP1, *Momordica charantia* ODP1 or *Glycine max* ODP1

T5 seed of *Arabidopsis* events C00536, 4445 and 4485 carrying pZBL120xKS336 (*Zea mays* ODP), pZBL120x

KS333 (*Momordica charantia* ODP1) and pZBL120xKS334 (*Glycine max* ODP1) transgenes, respectively, and WT seed derived from plants grown alongside each of the transgenic events were subjected to compositional analysis as described below. Seed weight was measured by determining the weight of 100 seed. This analysis was performed in triplicate.

##### Tissue preparation.

*Arabidopsis* seed (approximately 0.5 g in a 1/2x2" polycarbonate vial) was ground to a homogeneous paste in a GENO-GRINDER® (3x30 sec at 1400 strokes per minute, with a 15 sec interval between each round of agitation). After the second round of agitation the vials were removed and the *Arabidopsis* paste was scraped from the walls with a spatula prior to the last burst of agitation.

##### Determination of Protein Content:

Protein contents were estimated by combustion analysis on a Thermo FINNIGAN™ Flash 1112EA combustion analyzer running in the NCS mode (vanadium pentoxide was omitted) according to instructions of the manufacturer. Triplicate samples of the ground pastes, 4-8 mg, weighed to an accuracy of 0.001 mg on a METTLER-TOLEDO® MX5 micro balance, were used for analysis. Protein contents were calculated by multiplying % N, determined by the analyzer, by 6.25. Final protein contents were expressed on a % tissue weight basis.

##### Determination of Non-Structural Carbohydrate Content:

Sub-samples (30-35 mg) of the ground paste were weighed (to an accuracy of 0.1 mg) into 13x100 mm glass tubes; the tubes had TEFLON® lined screw-cap closures. Three replicates were prepared for each sample tested.

Lipid extraction was performed by adding 2 ml aliquots of heptane to each tube. The tubes were vortex mixed and placed into an ultrasonic bath (VWR Scientific Model 750D) filled with water heated to 60° C. The samples were sonicated at full-power (~360 W) for 15 min and were then centrifuged (5 minx1700 g). The supernatants were transferred to clean 13x100 mm glass tubes and the pellets were extracted 2 more times with heptane (2 ml, second extraction; 1 ml third extraction) with the supernatants from each extraction being pooled. After lipid extraction 1 ml acetone was added to the pellets and after vortex mixing, to fully disperse the material, they were taken to dryness in a Speedvac.

##### Non-Structural Carbohydrate Extraction and Analysis.

Two ml of 80% ethanol was added to the dried pellets from above. The samples were thoroughly vortex mixed until the plant material was fully dispersed in the solvent prior to sonication at 60° C. for 15 min. After centrifugation, 5 minx1700 g, the supernatants were decanted into clean 13x100 mm glass tubes. Two more extractions with 80% ethanol were performed and the supernatants from each were pooled. The extracted pellets were suspended in acetone and dried (as above). An internal standard  $\alpha$ -phenyl glucopyranoside (100  $\mu$ l of a 0.5000+/-0.0010 g/100 ml stock) was added to each extract prior to drying in a Speedvac. The extracts were maintained in a desiccator until further analysis.

The acetone dried powders from above were suspended in 0.9 ml MOPS (3-N[Morpholino]propane-sulfonic acid; 50 mM, 5 mM CaCl<sub>2</sub>, pH 7.0) buffer containing 100 U of heat-stable  $\alpha$ -amylase (from *Bacillus licheniformis*; Sigma A-4551). Samples were placed in a heat block (90° C.) for 75 min and were vortex mixed every 15 min. Samples were then allowed to cool to room temperature and 0.6 ml acetate buffer (285 mM, pH 4.5) containing 5 U amyloglucosidase (Roche 110 202 367 001) was added to each. Samples were incubated for 15-18 h at 55° C. in a water bath fitted with a reciprocating shaker; standards of soluble potato starch (Sigma S-2630) were included to ensure that starch digestion went to completion.

## 31

Post-digestion the released carbohydrates were extracted prior to analysis. Absolute ethanol (6 ml) was added to each tube and after vortex mixing the samples were sonicated for 15 min at 60° C. Samples were centrifuged (5 min×1700 g) and the supernatants were decanted into clean 13×100 mm glass tubes. The pellets were extracted 2 more times with 3 ml of 80% ethanol and the resulting supernatants were pooled. Internal standard (100  $\mu$ l  $\alpha$ -phenyl glucopyranoside, as above) was added to each sample prior to drying in a Speedvac.

## Sample Preparation and Analysis.

The dried samples from the soluble and starch extractions described above were solubilized in anhydrous pyridine (Sigma-Aldrich P57506) containing 30 mg/ml of hydroxylamine HCl (Sigma-Aldrich 159417). Samples were placed on an orbital shaker (300 rpm) overnight and were then heated for 1 hr (75° C.) with vigorous vortex mixing applied every 15 min. After cooling to room temperature, 1 ml hexamethyl-

## 32

isilazane (Sigma-Aldrich H-4875) and 100  $\mu$ l trifluoroacetic acid (Sigma-Aldrich T-6508) were added. The samples were vortex mixed and the precipitates were allowed to settle prior to transferring the supernatants to GC sample vials.

Samples were analyzed on an Agilent 6890 gas chromatograph fitted with a DB-17MS capillary column (15 m×0.32 mm×0.25  $\mu$ m film). Inlet and detector temperatures were both 275° C. After injection (2  $\mu$ l, 20:1 split) the initial column temperature (150° C.) was increased to 180° C. at a rate of 3° C./min and then at 25° C./min to a final temperature of 320° C. The final temperature was maintained for 10 min. The carrier gas was H<sub>2</sub> at a linear velocity of 51 cm/sec. Detection was by flame ionization. Data analysis was performed using Agilent ChemStation software. Each sugar was quantified relative to the internal standard and detector responses were applied for each individual carbohydrate (calculated from standards run with each set of samples). Final carbohydrate concentrations were expressed on a tissue weight basis.

TABLE 6

Composition Analysis of pZBL120xKS336, pZBL120xKS333 and pZBL120xKS334 Transgenic Seed and WT Control Seed					
Construct	Event ID	Oil (% NMR)	Protein %	Seed Weight ( $\mu$ g)	fructose ( $\mu$ g mg <sup>-1</sup> seed)
pZBL120xKS336	C00536	46.7	15.7	24	0.6
	WT	42	18.1	24	1
	$\Delta$ □TG/WT %	11.2	-13.3	0.0	-40.0
Construct	Event ID	glucose ( $\mu$ g mg <sup>-1</sup> seed)	sucrose ( $\mu$ g mg <sup>-1</sup> seed)	raffinose ( $\mu$ g mg <sup>-1</sup> seed)	stachyose ( $\mu$ g mg <sup>-1</sup> seed)
pZBL120xKS336	C00536	8.5	17.2	0.4	2.1
	WT	12.1	29.2	0.8	3.1
	$\Delta$ □TG/WT %	-29.8	-41.1	-50.0	-32.3
Construct	Event ID	Oil (% NMR)	Protein %	Seed Weight ( $\mu$ g)	fructose ( $\mu$ g mg <sup>-1</sup> seed)
pZBL120xKS333	4445	46	15	21.7	1
	WT	43.7	14.8	20.7	1.2
	$\Delta$ □TG/WT %	5.3	1.4	4.8	-16.7
Construct	Event ID	glucose ( $\mu$ g mg <sup>-1</sup> seed)	sucrose ( $\mu$ g mg <sup>-1</sup> seed)	raffinose ( $\mu$ g mg <sup>-1</sup> seed)	stachyose ( $\mu$ g mg <sup>-1</sup> seed)
pZBL120xKS333	4445	7.8	14.6	0.5	2
	WT	10.3	26.6	0.6	3.6
	$\Delta$ □TG/WT %	-24.3	-45.1	-16.7	-44.4
Construct	Event ID	Oil (% NMR)	Protein %	Seed Weight ( $\mu$ g)	fructose ( $\mu$ g mg <sup>-1</sup> seed)
pZBL120xKS334	4485	45.4	14.8	20.3	0.6
	WT	42.5	14.5	20.7	0.9
	$\Delta$ TG/WT %	6.8	2.1	-1.9	-33.3
Construct	Event ID	glucose ( $\mu$ g mg <sup>-1</sup> seed)	sucrose ( $\mu$ g mg <sup>-1</sup> seed)	raffinose ( $\mu$ g mg <sup>-1</sup> seed)	stachyose ( $\mu$ g mg <sup>-1</sup> seed)
pZBL120xKS334	4485	6.3	11.7	0.5	1.6
	WT	10.4	30.4	0.7	3.3
	$\Delta$ □TG/WT %	-39.4	-61.5	-28.6	-51.5

Table 6 shows that a reduction of soluble carbohydrates is consistently associated with presence of the pZBL120×KS333, 334 and 336 transgenes. There is no consistent change in protein content or seed weight that can be attributed to the pZBL120×KS333, 334 and 336 transgenes.

Example 7

Germination Assays of *Arabidopsis* Seed Transformed with *Zea mays* ODP1, *Momordica charantia* ODP1 or *Glycine max* ODP1

T5 seed of *Arabidopsis* events C00536, 4445 and 4485 carrying pZBL120×KS336 (*Zea mays* ODP1), pZBL120×KS333 (*Momordica charantia* ODP1) and pZBL120×KS334 (*Glycine max* ODP1) transgenes, respectively, were subjected to germination assays on standard *Arabidopsis* growth media (see above) containing either 10 g L<sup>-1</sup> sucrose or equimolar amounts of sorbitol (5.3 g L<sup>-1</sup>). Seeds were surface-sterilized and homogeneous dispersion of the seed on the medium was facilitated by mixing the aqueous seed suspension with an equal volume of melted plant growth medium containing the either sucrose or sorbitol. Plates were incubated under standard conditions (22° C., 16 h light/8 h dark, 100 μE m<sup>-2</sup>s<sup>-1</sup>) and germination rate and seedling phenotype was scored 14 days after plating (Table 7).

TABLE 7

Germination Assays for pZBL120×KS336, pZBL120×KS333 and pZBL120×KS334 Transgenic Seeds					
Line ID	Media Type	Total Seed (#)	Altered Seedling Morphology (#)	No Germination (#)	Healthy Seedlings (#)
C00536	sucrose	93	69	2	22
C00536	sucrose	84	50	3	31
C00536	sucrose	90	73	3	14
C00536	sorbitol	95	6	89	0
C00536	sorbitol	112	24	88	0
C00536	sorbitol	100	49	51	0
4445	sucrose	82	24	22	36
4445	sucrose	63	24	7	32
4445	sucrose	94	36	12	46
4445	sorbitol	106	70	36	0
4445	sorbitol	119	77	42	0
4445	sorbitol	106	97	9	0
4485	sucrose	98	50	48	0
4485	sucrose	109	37	70	2
4485	sucrose	129	80	39	10
4485	sorbitol	131	24	107	0
4485	sorbitol	128	25	103	0
4485	sorbitol	127	23	102	2

Line ID	Media Type	Altered Seedling Morphology (%)	No Germination (%)	Healthy Seedlings (%)
C00536	sucrose	74.2	2.2	23.7
C00536	sucrose	59.5	3.6	36.9
C00536	sucrose	81.1	3.3	15.6
	AVG	71.6	3.0	25.4
C00536	sorbitol	6.3	93.7	0.0
C00536	sorbitol	21.4	78.6	0.0
C00536	sorbitol	49.0	51.0	0.0
	AVG	25.6	74.4	0.0
4445	sucrose	29.3	26.8	43.9

TABLE 7-continued

Germination Assays for pZBL120×KS336, pZBL120×KS333 and pZBL120×KS334 Transgenic Seeds				
4445	sucrose	38.1	11.1	50.8
4445	sucrose	38.3	12.8	48.9
	AVG	35.2	16.9	47.9
4445	sorbitol	66.0	34.0	0.0
4445	sorbitol	64.7	35.3	0.0
4445	sorbitol	91.5	8.5	0.0
	AVG	74.1	25.9	0.0
4485	sucrose	51.0	49.0	0.0
4485	sucrose	33.9	64.2	1.8
4485	sucrose	62.0	30.2	7.8
	AVG	49.0	47.8	3.2
4485	sorbitol	18.3	81.7	0.0
4485	sorbitol	19.5	80.5	0.0
4485	sorbitol	18.1	80.3	1.6
	AVG	18.7	80.8	0.5

It is evident that germination and/or seedling development is significantly affected in all events analyzed. Germination is improved in the presence of sucrose; however, in events carrying pZBL120×KS336 and pZBL120×KS334 the seed germinating on sucrose containing media gave rise to seedlings with altered morphology, namely the presence of leaf structures that fail to become green and which resemble non-photosynthetic cotyledon tissue.

Total fatty acid (FA) composition and content of seedling tissue of C00536, 4485 and WT seedlings were measured 14 days after plating on media containing 10 g L<sup>-1</sup> sucrose. Briefly, seedling tissue was frozen on dry ice or by incubation in a -80° C. freezer for two h followed by lyophilization for 48 h.

Dried seedling tissue was ground to a fine powder using a GENOGRINDER® vial (1/2"×2" polycarbonate) and a steel ball (SPEX Centriprep (Metuchen, N.J., U.S.A.)). Grinding time was 30 sec at 1450 oscillations per min. Ten mg of tissue were weighed into Eppendorf tubes. The tissue was extracted using 100 μL heptane at room temperature under continuous shaking for 2 h. Heptane extracts were cleared by centrifugation and 25 μL of extract was derivatized to fatty acid methyl esters as follows. One mL of a 25% sodium methoxide stock solution was added to 24 mL of HPLC grade methanol. Sodium methoxide was stored under an inert gas.

Five μL of a 17:0 TAG (Nu-Chek Prep, Elysian, Minn., USA) stock solution (10 mg/mL) was combined with 25 μL of heptane tissue extract in a glass culture tube and 500 μL of 1% sodium methoxide was added. Samples were derivatized in a water bath at 50° C. for 15 min. Samples were allowed to cool to RT and 1 mL of 1M NaCl was added followed by brief mixing. FAMES were extracted into 1 mL of heptane and 4 μL sample were quantitated by GC analysis (Table 8).

TABLE 8

Fatty Acid Composition and Total Fatty Acid Content of Seedling Tissue of WT Plants and pZBL120xKS334 and pZBL120xKS336 Transgenic Plants Grown on Sucrose-Containing Media										
Event ID	% Total FA									Total FA (% DW)
	16:0	18:0	18:1	18:2	18:3	20:0	20:1	20:2	22:0	
WT	13.5	10.0	42.3	15.2	15.6	1.0	0.9	0.0	1.5	4.3
4485	11.8	2.7	13.2	26.6	26.7	2.5	13.5	2.2	0.7	18.6
C00536	7.9	2.3	15.0	17.7	32.4	3.2	18.2	2.0	1.2	21.9

Table 8 demonstrates that seedling tissue of transgenic lines carrying pZBL120xKS334 and pZBL120xKS336 transgenes showed increased fatty acid content when compared to WT seedlings. Moreover, the fatty acid profile of transgenic seedling tissue is similar to that of *Arabidopsis* WT seed in that it contains significant levels (>15%) of C20 fatty acids.

In summary, use of a strong heterologous seed storage protein promoter (soybean  $\beta$ -conglycinin promoter) for expression in *Arabidopsis* of ODP1 genes from a diverse range of plant species belonging to the families of Leguminosae, Cucurbitaceae and Poaceae, resulted in increased seed storage lipid accumulation at the expense of soluble carbohydrates. However, seed germination and seedling establishment was negatively affected in transgenic lines expressing ODP1 genes under control of a strong heterologous seed storage protein promoter.

#### Example 8

##### Construction of *Arabidopsis* Expression Vector pKR1223 for Expression of *Zea mays* ODP Under Control of the Seed-Specific, Low Strength *Arabidopsis* Sucrose Synthase Promoter

The present example describes the synthesis of *Arabidopsis* expression vector pKR1223 which allows for expression of the *Zea mays* ODP gene under control of the promoter of an *Arabidopsis* sucrose synthase gene (At5g49190). Additionally, vector pKR1223 provides seed-specific expression of the DsRed gene in order to visualize positive transformants as well as constitutive expression of the npt gene for selection on kanamycin.

Plasmid pKR132 (SEQ ID NO:16) which is described in PCT Publication No. WO 2004/071467 (the contents of which are incorporated by reference), was digested with BamHI/SalI and the fragment containing the soy albumin promoter was cloned into the BamHI/XhoI fragment of the pCR-Blunt® cloning vector (Invitrogen Corporation) to produce the starting vector pKR627 (SEQ ID NO:17).

Plasmid KS294 (SEQ ID NO:18) contains a NotI site flanked by the SCP1 promoter and the phaseolin transcription terminator (SCP1Pro::NotI::PhasTerm). The SCP1 promoter is a synthetic constitutive promoter comprising a portion of the CaMV 35S promoter (Odell et al. (1985) Nature 313:810-812) and the Rsyn7-Syn II Core synthetic consensus promoter (U.S. Pat. Nos. 6,072,050 and 6,555,673, the contents of which are incorporated by reference). See also, for example, US20030226166, Table 13 (the contents of which are incorporated by reference). Downstream of this element is the Tobacco Mosaic Virus (TMV) omega 5'-UTR translational enhancer element (Gallie et al. (1992) Nucleic Acid Research 20:4631-4638), followed by the NotI site and the 3' transcription termination region of the phaseolin gene (Doyle

et al., (1986) *J. Biol. Chem.* 261:9228-9238). The XbaI fragment of KS294 (SEQ ID NO:18), containing the SCP1Pro::NotI::PhasTerm cassette, was cloned into the XbaI site of pKR627 (SEQ ID NO:17) to produce pKR1142 (SEQ ID NO:19).

The BamHI fragment of KS334 (SEQ ID NO:14; Example 1), containing the Kti3Pro:DsRed:Kti3Term cassette, was cloned into the BamHI site of pKR278 (SEQ ID NO:20), which was previously described in U.S. Patent Publication No. US20080095915 (the contents of which are incorporated by reference), to produce vector pKR1141 (SEQ ID NO:20).

Genomic DNA was isolated from 3 week-old wild-type *Arabidopsis* col-0 seedlings using the DNEASY® Plant Mini Kit (Qiagen, Valencia, Calif.) and following the manufacturer's protocol. An *Arabidopsis* Sucrose Synthase ("AtSuSy"; "AtSUS2") promoter derived from gene At5g49190 was PCR-amplified from *Arabidopsis* genomic DNA using oligonucleotides SuSy-5 (SEQ ID NO:21) and SuSy-3 (SEQ ID NO:22) with the PHUSION™ High-Fidelity DNA Polymerase (Cat. No. F553S, Finnzymes Oy, Finland), following the manufacturer's protocol. The resulting DNA fragment was cloned into the pCR®-BLUNT® cloning vector using the ZERO BLUNT® PCR Cloning Kit (Invitrogen Corporation), following the manufacturer's protocol, to produce pLF122 (SEQ ID NO:23).

The BamHI/NotI fragment of pLF122 (SEQ ID NO:23), containing the AtSuSy promoter, was cloned into the BamHI/NotI fragment of pKR1142 (SEQ ID NO:19), containing the phaseolin terminator, to produce pKR1155 (SEQ ID NO:24).

The Asp718/BsiWI fragment of pKR1155 (SEQ ID NO:24), containing the AtSuSy promoter, was cloned into the BsiWI site of pKR1141 (SEQ ID NO:20), to produce pKR1158 (SEQ ID NO:25).

The NotI fragment of KS336 (SEQ ID NO:6; Example 1), containing the corn ODP, was cloned into the NotI site of pKR1158 (SEQ ID NO:25), to produce pKR1167 (SEQ ID NO:26).

The AscI fragment of pKR1167 (SEQ ID NO:26), containing the corn ODP gene, was cloned into the AscI fragment of pKR92 (SEQ ID NO:27) which was previously described in WO2007/061845 (published on May 31, 2007, the contents of which are herein incorporated by reference) to produce pKR1223 (SEQ ID NO:28).

#### Example 9

##### Construction of *Arabidopsis* Expression Vector pKR1220 for Expression of the Corn ODP Under Control of the Seed-Specific, Medium-Strength Soy Annexin Promoter

The present example describes the synthesis of *Arabidopsis* expression vector pKR1220 which allows for seed-specific expression of the corn ODP gene under control of the soy

annexin promoter. Additionally, vector pKR1220 provides seed-specific expression of the DsRed gene in order to visualize positive transformants and constitutive expression of the npt gene for selection on kanamycin.

The BsiWI fragment of pKR268 (SEQ ID NO:29; which is described in PCT Publication No. WO 04/071467, the contents of which are herein incorporated by reference), containing the AnnexinPro::NotI::BD30Term cassette, was cloned into the BsiWI site of pKR1141 (SEQ ID NO:20) to give pKR1143 (SEQ ID NO:30).

The NotI fragment of KS336 (SEQ ID NO:6), containing the corn ODP1 gene, was cloned into the NotI site of pKR1143 (SEQ ID NO:30), to produce pKR1147 (SEQ ID NO:31).

The AscI fragment of pKR1147 (SEQ ID NO:31), containing the corn ODP1 gene, was cloned into the AscI fragment of pKR92 (SEQ ID NO:27) to produce pKR1220 (SEQ ID NO:32).

#### Example 10

##### Construction of *Arabidopsis* Expression Vector pKR1221 for Expression of the Corn ODP Under Control of the Constitutive, Medium Strength SCP1 Promoter

The present example describes the synthesis of *Arabidopsis* expression vector pKR1221 which allows for constitutive expression of the corn ODP1 gene under control of the SCP1 promoter. Additionally, vector pKR1221 provides seed-specific expression of the DSred gene in order to visualize positive transformants and constitutive expression of the npt gene for selection on kanamycin.

The Asp718/BsiWI fragment of pKR1142 (SEQ ID NO:19), containing the SCP1Pro::NotI::PhasTerm cassette, was cloned into the BsiWI site of pKR1141 (SEQ ID NO:20), to produce pKR1144 (SEQ ID NO:33).

The NotI fragment of KS336 (SEQ ID NO:6), containing the corn ODP1, was cloned into the NotI site of pKR1144 (SEQ ID NO:33), to produce pKR1149 (SEQ ID NO:34).

The AscI fragment of pKR1149 (SEQ ID NO:34), containing the corn ODP1 gene, was cloned into the AscI fragment of pKR92 (SEQ ID NO:27) to produce pKR1221 (SEQ ID NO:35).

#### Example 11

##### Generation and Analysis of T2 Seed of *Arabidopsis* Lines Transformed with Corn ODP Under Control of the SCP1, Annexin or Sucrose Synthase Promoters

Plasmid DNA of pKR1220, pKR1221 and pKR1223 was introduced into *Agrobacterium tumefaciens* NTL4 (Luo et al, *Molecular Plant-Microbe Interactions* (2001) 14(1):98-103) by electroporation. Briefly, 1 µg plasmid DNA was mixed with 100 µL of electro-competent cells on ice. The cell suspension was transferred to a 100 µL electroporation cuvette (1 mm gap width) and electroporated using a BIORAD electroporator set to 1 kV, 400Ω and 25 µF. Cells were transferred to 1 mL LB medium and incubated for 2 h at 30° C. Cells were plated onto LB medium containing 50 µg/mL kanamycin. Plates were incubated at 30° C. for 60 h. Recombinant *Agrobacterium* cultures (500 mL LB, 50 µg/mL kanamycin) were inoculated from single colonies of transformed *Agrobacterium* cells and grown at 30° C. for 60 h. Cells were harvested by centrifugation (5000×g, 10 min) and resuspended in 1 L of 5% (W/V) sucrose containing 0.05% (V/V) Silwet. *Arabi-*

*dopsis* plants were grown in soil at a density of 30 plants per 100 cm<sup>2</sup> pot in METRO-MIX® 360 soil mixture for 4 weeks (22° C., 16 h light/8 h dark, 100 µE m<sup>-2</sup>s<sup>-1</sup>). Plants were repeatedly dipped into the *Agrobacterium* suspension harboring the relevant binary vector and kept in a dark, high humidity environment for 24 h. Plants were grown for three to four weeks under standard plant growth conditions described above and plant material was harvested and dried for one week at ambient temperatures in paper bags. Seeds were harvested using a 0.425 mm mesh brass sieve.

Cleaned *Arabidopsis* seeds (2 grams, corresponding to about 100,000 seeds) were sterilized by washes in 45 mL of 80% ethanol, 0.01% TRITON® X-100, followed by 45 mL of 30% (V/V) household bleach in water, 0.01% TRITON® X-100 and finally by repeated rinsing in sterile water. Aliquots of 20,000 seeds were transferred to square plates (20×20 cm) containing 150 mL of sterile plant growth medium comprised of 0.5×MS salts, 0.53% (W/V) sorbitol, 0.05 MES/KOH (pH 5.8), 200 µg/mL TIMENTIN®, and 50 µg/mL kanamycin solidified with 10 g/L agar. Homogeneous dispersion of the seed on the medium was facilitated by mixing the aqueous seed suspension with an equal volume of melted plant growth medium. Plates were incubated under standard growth conditions for ten days. Kanamycin-resistant seedlings were transferred to plant growth medium without selective agent and grown for one week before transfer to soil. Plants were grown to maturity and T2 seeds were harvested and plated on selective media containing kanamycin. Approximately 100 events were generated in this manner. Wild-type control plants were grown in the same flat containing transgenic T1 plants. T2 seeds were harvested and oil content was measured by NMR (Tables 9 and 10).

TABLE 9

Data from Germination Assays for T2 Seed of pKR1220, pKR1221 and pKR1223 Transgenics on Selective Medium Containing Kanamycin and Sorbitol

Event ID	pKR	Total Seed (#)	Transgenic Seed (#)	ASM* (#)	Kan <sup>S</sup> (#)	No Germination (#)	Healthy Seed-Lings (#)	Δ Oil % points
35634	1220	122	110	11	12	31	68	2.6
36062	1220	134	127	25	7	85	17	2.4
35637	1220	147	133	16	14	100	17	2.4
36066	1220	143	123	22	20	59	42	2
35636	1220	116	105	19	11	62	24	1.7
36059	1220	101	85	14	16	52	19	1.6
36104	1221	104	104	6	0	96	2	4.7
36078	1221	83	66	0	17	66	0	3
36087	1221	93	89	0	4	89	0	2
36090	1221	103	103	1	0	98	4	1.9
36101	1221	134	126	0	8	126	0	1.7
36122	1221	108	92	0	16	92	0	1.7
36162	1223	92	83	8	9	20	55	5.3
36210	1223	112	111	2	1	21	88	4.4
36151	1223	144	142	66	2	40	36	3.6
36194	1223	94	91	14	3	11	66	3.4
36157	1223	101	77	14	24	10	53	3.4
36181	1223	160	149	15	11	88	46	3.3
36199	1223	103	95	17	8	12	66	3.2
36208	1223	119	110	22	9	20	68	3.1
36161	1223	134	120	19	14	33	68	3
36200	1223	144	140	0	4	101	39	2.8
36154	1223	110	99	10	11	7	82	2.7
36209	1223	109	106	10	3	31	65	2.6
36179	1223	172	147	10	25	68	69	2.6
36180	1223	162	149	16	13	51	82	2.6
36213	1223	146	127	22	19	57	48	2.4
36206	1223	86	79	17	7	0	62	2.2

\*ASM denotes Altered Seedling Morphology

TABLE 10

Results from Germination Assays for T2 Seed of pKR1220, pKR1221 and pKR1223 Transgenics on Selective Medium Containing Kanamycin and Sorbitol					
Event ID	pKR	% ASM*	% No Germination	% Healthy Seedlings	Δ Oil % Points
35634	1220	10.0	28.2	61.8	2.6
36062	1220	19.7	66.9	13.4	2.4
35637	1220	12.0	75.2	12.8	2.4
36066	1220	17.9	48.0	34.1	2.0
35636	1220	18.1	59.0	22.9	1.7
36059	1220	16.5	61.2	22.4	1.6
	AVG	15.7	56.4	27.9	2.1
36104	1221	5.8	92.3	1.9	4.7
36078	1221	0.0	100.0	0.0	3.0
36087	1221	0.0	100.0	0.0	2.0
36090	1221	1.0	95.1	3.9	1.9
36101	1221	0.0	100.0	0.0	1.7
36122	1221	0.0	100.0	0.0	1.7
	AVG	1.1	97.9	1.0	2.5
36162	1223	9.6	24.1	66.3	5.3
36210	1223	1.8	18.9	79.3	4.4
36151	1223	46.5	28.2	25.4	3.6
36194	1223	15.4	12.1	72.5	3.4
36157	1223	18.2	13.0	68.8	3.4
36181	1223	10.1	59.1	30.9	3.3
36199	1223	17.9	12.6	69.5	3.2
36208	1223	20.0	18.2	61.8	3.1
36161	1223	15.8	27.5	56.7	3.0
36200	1223	0.0	72.1	27.9	2.8
36154	1223	10.1	7.1	82.8	2.7
36209	1223	9.4	29.2	61.3	2.6
36179	1223	6.8	46.3	46.9	2.6
36180	1223	10.7	34.2	55.0	2.6
36213	1223	17.3	44.9	37.8	2.4
36206	1223	21.5	0.0	78.5	2.2
	AVG	14.4	28.0	57.6	3.2

\*\*"ASM" denotes Altered Seedling Morphology

Example 12

Analysis of T3 and T4 Seed of *Arabidopsis* Plants Transformed with *Zea mays* ODP Under Control of the *Arabidopsis* Sucrose Synthase Promoter

T2 seeds of pKR1223 transformation events 36162, 36180 and 36181 were germinated on selective media containing kanamycin. Twenty-four kanamycin-resistant seedlings were planted in soil along side twelve untransformed WT *Arabidopsis* plants. Plants were grown to maturity and T3 seed samples were harvested from individual T2 plants. A bulk seed sample was generated from all WT plants in a given flat. Oil content was measured by NMR (Table 11).

TABLE 11

Oil Content of T3 Seed of pKR1223 Transgenics		
Event	Plant #	% oil
36162	1	44.6
36162	2	44.5
36162	3	44.4
36162	4	44.3
36162	5	44.3
36162	6	44.2
36162	7	44.2
36162	8	43.9
36162	9	43.8
36162	10	43.7
36162	11	43.7
36162	12	43.7

TABLE 11-continued

Oil Content of T3 Seed of pKR1223 Transgenics		
Event	Plant #	% oil
36162	13	43.7
36162	14	43.7
36162	15	43.6
36162	16	43.5
36162	17	43.5
36162	18	43.5
36162	19	43.4
36162	20	43.0
36162	21	42.8
36162	22	42.2
36162	23	41.8
36162	24	36.4
36162	AVG	43.4
36162	Exp.	41.8
36180	1	44.5
36180	2	44.3
36180	3	43.8
36180	4	43.8
36180	5	43.7
36180	6	43.6
36180	7	43.6
36180	8	43.6
36180	9	43.5
36180	10	43.4
36180	11	43.3
36180	12	43.3
36180	13	43.3
36180	14	43.3
36180	15	43.2
36180	16	43.2
36180	17	43.1
36180	18	43.1
36180	19	42.9
36180	20	42.9
36180	21	42.8
36180	22	42.8
36180	23	42.7
36180	24	42.6
36180	AVG	43.3
36180	WT in	
36180	Exp.	41.9
36181	1	47.2
36181	2	46.3
36181	3	46.2
36181	4	46.1
36181	5	45.9
36181	6	45.7
36181	7	45.4
36181	8	45.0
36181	9	45.0
36181	10	45.0
36181	11	45.0
36181	12	44.9
36181	13	44.9
36181	14	44.8
36181	15	44.7
36181	16	44.6
36181	17	44.5
36181	18	44.4
36181	19	44.4
36181	20	43.8
36181	21	43.8
36181	22	43.6
36181	23	43.3
36181	24	42.6
36181	AVG	44.9
36181	WT in	
36181	Exp.	41.9

41

Transgenic T3 seed selections of events 36180 and 36162 that no longer segregated for the DsRed marker gene were identified by visual inspection using a suitable light source. These T3 selections that were homozygous for the pKR1223 transgene were subjected to germination assays on plant growth media containing sucrose or sorbitol as described above (Table 12).

TABLE 12

Germination Assays for T3 Seed of pKR1223 Transgenics					
Event	Media Type	Total Seed (#)	ASM * (#)	No Germination (#)	Healthy Seedlings (#)
36180	sucrose	83	0	0	83
36180	sucrose	111	0	0	111
36180	sucrose	110	0	0	110
36180	sorbitol	121	0	0	121
36180	sorbitol	128	0	0	128
36180	sorbitol	118	0	0	118
36162	sucrose	88	0	0	88
36162	sucrose	111	1	1	109
36162	sucrose	90	0	0	90
36162	sorbitol	97	0	0	97
36162	sorbitol	103	0	0	103
36162	sorbitol	107	2	0	105

Event	Media Type	ASM * (%)	No Germination (%)	Healthy Seedlings (%)
36180	sucrose	0.0	0.0	100.0
36180	sucrose	0.0	0.0	100.0
36180	sucrose	0.0	0.0	100.0
36180	sucrose	AVG	0.0	100.0
36180	sorbitol	0.0	0.0	100.0
36180	sorbitol	0.0	0.0	100.0
36180	sorbitol	0.0	0.0	100.0
36180	sorbitol	AVG	0.0	100.0
36162	sucrose	0.0	0.0	100.0
36162	sucrose	0.9	0.9	98.2
36162	sucrose	0.0	0.0	100.0
36162	sucrose	AVG	0.3	99.4
36162	sorbitol	0.0	0.0	100.0
36162	sorbitol	0.0	0.0	100.0
36162	sorbitol	1.9	0.0	98.1
36162	sorbitol	AVG	0.6	99.4

\* "ASM" denotes Altered Seedling Morphology

Transgenic T3 seed selections of events 36180 and 36162 that no longer segregated for the DsRed marker gene were identified by visual inspection using a suitable light source. In case of event 36181 no T3 seed selections could be identified that did not segregate for the DS red marker in a total of 24 progeny seed samples derived from 24 kanamycin-resistant T2 plants. Moreover, when T3 seed were plated on selective agarose media, 25% of seed failed to germinate and 25% of the seedlings were sensitive to kanamycin. It is concluded that the transgene insertion in event 36181 can only be maintained in the heterozygous state. The homozygous nature of T3 seed selections of events 36180 and 36162 suggests that the seed phenotype of event 36181 is related to the transgene insertion site and not the transgene itself. It is believed that a gene that is important for development of viable seed was disrupted by the transgene insertion.

T3 seed selections of events 36180 and 36162 that were homozygous for the transgene insertion and T3 seed selections of event 36181 that were heterozygous for the transgene insertion were germinated on selective media containing kanamycin. Three flats were planted for every transgenic event as follows: 24 seedlings were planted in each flat next to 12 WT seedlings at identical developmental stage. Plants

42

were grown to maturity for approximately eight weeks and seed were harvested in bulk from all transgenic and WT plants in a given flat. Oil content of seed was measured by NMR as described in Example 1. Results are summarized in Table 13. In all three events presence of the pKR1223-derived transgene leads to an increase in oil content that ranges between 0.7 and 2.2% points (1.6-5.4%).

TABLE 13

Oil Content of T4 Seed of pKR1223 Transgenics					
Flat ID	Event ID	Oil (%)	Δ Oil (%) Points)	Δ Oil (%)	
15	A	36181	42.8	2.2	5.4
		WT	40.6		
	B	36181	43.5	2.1	5.2
		WT	41.4		
	C	36181	40.8	1.5	4.0
		WT	39.2		
		AVG	40.5	2.0	4.9
20	A	36180	44.5	1.8	4.2
		WT	42.7		
	B	36180	43.6	1.9	4.6
		WT	41.7		
	C	36180	43.2	1.2	2.8
		WT	42.0		
		AVG	43.0	1.6	3.9
25	A	36162	43.3	1.4	3.4
		WT	41.9		
	B	36162	43.6	0.7	1.6
		WT	42.9		
	C	36162	43.8	1.0	2.4
		WT	42.7		
		AVG	43.2	1.0	2.5

T4 seed of events 36162 and 36180 were subjected to compositional analysis as described in Example 6.

TABLE 14

Composition of pKR1223 Transgenic T4 Seed and WT Control Seed					
Event	Oil (% NMR)	Protein (%)	Seed Weight (μg)	Fructose (μg mg <sup>-1</sup> seed)	
36162	43.3	14.94	20.33	2.13	
WT	41.9	15.05	19	2.39	
Δ □TG/WT %	3.3	-0.7	7.0	-10.9	

Event	Glucose (μg mg <sup>-1</sup> seed)	Sucrose (μg mg <sup>-1</sup> seed)	Raffinose (μg mg <sup>-1</sup> seed)	Stachyose (μg mg <sup>-1</sup> seed)
36162	4.82	11.32	0.56	1.52
WT	5.17	14.28	0.64	1.58
ΔTG/WT %	-6.8	-20.7	-12.5	-3.8

Event	Oil (% NMR)	Protein (%)	Seed Weight (μg)	Fructose (μg mg <sup>-1</sup> seed)
36180	43.6	15.17	21	2.07
WT	41.7	15.16	21	2.45
Δ TG/WT %	4.6	0.1	0.0	-15.5

Event	Glucose (μg mg <sup>-1</sup> seed)	Sucrose (μg mg <sup>-1</sup> seed)	Raffinose (μg mg <sup>-1</sup> seed)	Stachyose (μg mg <sup>-1</sup> seed)
36180	4.49	11.14	0.5	1.46
WT	4.97	14.08	0.57	1.45

TABLE 14-continued

Composition of pKR1223 Transgenic T4 Seed and WT Control Seed				
$\Delta$ □TG/WT %	-9.7	-20.9	-12.3	0.7

A reduction of soluble carbohydrates (mainly sucrose) was consistently associated with the presence of the pKR1223 transgene in events 36162 and 36180. There was no consistent change in protein content or seed weight that can be attributed to the presence of the transgene.

In summary, use of a promoter of the *Arabidopsis* sucrose synthase (SUS2) gene (At5g49190) for expression of maize ODP1 resulted in increased seed storage lipid accumulation at the expense of soluble carbohydrates. Seed germination and seedling establishment was not affected.

#### Example 13

##### Identification of Seed Specific Promoters to Drive ODP1 Expression in Cruciferous Oilseed Plants

The sucrose synthase gene family and the role of specific gene family members during seed development, specifically the mobilization of sucrose for seed storage compound biosynthesis, has been described (Ruuska S A, Girke T, Benning C and Ohlrogge J B (2002) Contrapuntal networks of gene expression during *Arabidopsis* seed filling. *Plant Cell* 14: 1191-1206; Baud S, Vaultier M-N and Rochat C (2004) Structure and expression profile of the sucrose synthase multigene family in *Arabidopsis*. *J Exp Bot* 55: 397-409; and Baud S and Graham I A (2006) A spatiotemporal analysis of enzymatic activities associated with carbon metabolism in wild-type and mutant embryos of *Arabidopsis* using in situ histochemistry. *Plant J* 46: 155-169). The current invention describes the unexpected utility of a promoter sequence of a specific gene family member, At5g49190, to direct expression of heterologous ODP1 genes in a manner that allows for increased accumulation of oil during seed development of cruciferous oil seed without affecting germination and seedling establishment of the resulting seed. At5g49190 is expressed during seedling development in synchrony with accumulation of oil and protein (supra). Genes homologous to At5g49190 can be identified in other plant species based on sequence similarity to the At5g49190 gene product and expression pattern of the homolog during seed development. One skilled in the art will recognize that promoter sequences of these genes will have utility for expression of ODP1 genes for increased oil biosynthesis in cruciferous oil seed which is accompanied by unaltered seed germination and seedling establishment.

#### Example 14

##### Identification of Canola Promoters to Drive ODP1 Expression in Cruciferous Oilseed Plants

Public EST and genomic sequence collections of Canola were searched with the deduced amino acid sequence of At5g49190 (AtSUS2). Several ESTs and genomic sequences were identified and assembled into a single contiguous sequence that represents a transcript model of the canola homolog of At5g49190. The nucleotide and deduced amino acid sequence of the canola SUS2 homolog transcript model are set forth as SEQ ID NO:44 and SEQ ID NO:45, respectively.

Primers a (SEQ ID NO:46) b (SEQ ID NO:47) c (SEQ ID NO:48) and d (SEQ ID NO:49) were used in genome walking experiments according to manufacturer instructions (Clontech, CA, USA). Briefly genomic DNA of Pioneer Hi-Bred International, Inc., spring canola variety NS1822BC was isolated using standard protocols and digested with PvuII or DraI. After adaptor ligation PCR PvuII and DraI-digested genomic DNA was used as template in PCR reactions with Primer a (SEQ ID:46) and Primer c (SEQ ID NO:48), respectively. PCR products generated with primers a (SEQ ID NO:46) and c (SEQ ID NO:48) were amplified with primers b (SEQ ID NO:47) and d (SEQ ID NO:49), respectively. In both rounds of PCR experiments adaptor specific primers were used with primers a-d. Use of primers a and b generated PCR products of 2.1 kb. Primers c and d generated PCR products of 0.7 kb. These PCR products were cloned using the PCR blunt cloning system (Invitrogen, CA, USA) and sequenced.

SEQ ID NO:50 (PvuII rapa cons) is genomic sequence of canola variety NS1822BC that was generated with primers a and b. It is comprised of 312 bp of a canola SUS2 homolog and 1924 bp of sequence upstream of the inferred start codon of the SUS2 gene. This 1924 bp sequence (including the 5' untranslated region) is designated the BnSUS2-2 promoter (SEQ ID NO:73).

SEQ ID NO:51 (1,6 DraI gene cons) is genomic sequence of canola variety NS1822BC that was generated with primers c and d. It is comprised of 37 bp of a canola SUS2 gene and 586 bp of sequence upstream of the inferred start codon of the SUS2 gene. This 586 bp sequence (including the 5' untranslated region) is designated the BnSUS2-1 promoter (SEQ ID NO:72).

Plasmid DNA of clone #6 containing 1,6 DraI gene cons (SEQ ID NO:51) was used in a PCR reaction with primers SA188 (SEQ ID NO:52) and SA189 (SEQ ID NO:53) using PHUSION™ DNA polymerase (New England Biolabs, Inc.). Plasmid DNA of clone #45 containing PvuII rapa cons (SEQ ID NO:50) was used in a PCR reaction with primers SA190 (SEQ ID NO:54) and SA191 (SEQ ID NO:55). PCR products from both reactions were cloned into PCR blunt (Invitrogen, CA, USA) according to manufacturer instructions and sequenced. BN SUS2 prom1/PCR blunt is derived from 1,6 DraI gene cons (SEQ ID NO:51). It's sequence is set forth as SEQ ID NO:56. BN SUS2 prom2/PCR blunt is derived from PvuII rapa cons (SEQ ID NO:50). It's sequence is set forth as SEQ ID NO:57.

BN SUS2 prom1/PCR blunt (SEQ ID NO:56) was linearized with XbaI and NotI and ligated with a NotI-XbaI fragment from KS332 (SEQ ID NO:3) containing Phas terminator and Kti promoter DS red gene and Kti terminator cassette to give KS427 (SEQ ID NO:58). KS427 (SEQ ID NO:58) was linearized with NotI. A delta-6 desaturase gene of *Mortierella alpina* was excised from KS130 (SEQ ID NO:59) using NotI and ligated to NotI linearized KS427 (SEQ ID NO:58) to give KS432 (SEQ ID NO:60). Expression cassettes for DSred and delta-6 desaturase genes were excised as a single DNA fragment by digestion with AscI and inserted into AscI linearized pKR92 (SEQ ID NO:27) to give ARALO80 (SEQ ID NO:61). The ARALO80 vector contains the following expression unit: BnSUS2-1 promoter::M. alpina delta-6 desaturase::phaseolin terminator.

Prior to this KS130 (SEQ ID NO:59) was constructed as follows: Plasmid DNA of CGR-5, which is described in U.S. Pat. No. 5,968,809, was used in a PCR reaction with primers D6 fwd (SEQ ID NO:62) and D6 rev (SEQ ID NO:63). The PCR product was digested with NotI and ligated to NotI-linearized and de-phosphorylated KS119 vector (SEQ ID

NO:64) to give KS130 (SEQ ID NO:59). Vector KS119 (SEQ ID NO:64) is described in International Publication No. WO2004071467.

The maize ODP1 gene was excised from KS336 (SEQ ID NO:6) using NotI and ligated to NotI linearized KS427 (SEQ ID NO:58) to give KS430 (SEQ ID NO:65). Expression cassettes for DSred and maize ODP1 genes were excised as a single fragment by digestion with AscI and inserted into AscI linearized pKR92 (SEQ ID NO:27) to give ARALO78 (SEQ ID NO:66). The ARALO78 vector contains the following expression unit: BnSUS2-1 promoter::ZM-ODP1::phaseolin terminator.

BN SUS2 pro2/PCR blunt (SEQ ID NO:57) was linearized with XbaI and NotI and ligated with a NotI-XbaI fragment from KS332 (SEQ ID NO:3) containing Phas terminator and Kti promoter DS red gene and Kti terminator cassette to give KS428 (SEQ ID NO:67). KS428 (SEQ ID NO:67) was linearized with NotI. The delta-6 desaturase gene was excised from KS130 (SEQ ID NO:59) using NotI and ligated to NotI-linearized KS428 (SEQ ID NO:67) to give KS429 (SEQ ID NO:68). Expression cassettes for DSred and delta-6 desaturase genes were excised as a single DNA fragment by digestion with AscI and inserted into AscI linearized pKR92 (SEQ ID NO:27) to give ARALO77 (SEQ ID NO:69). The ARALO77 vector contains the following expression unit: BnSUS2-2 promoter::*M. alpina* delta-6 desaturase::phaseolin terminator.

The maize ODP1 gene was excised from KS336 (SEQ ID NO:6) using NotI and ligated to NotI-linearized KS428 (SEQ ID NO:67) to give KS431 (SEQ ID NO:70). Expression cassettes for DSred and maize ODP1 genes were excised by digestion with AscI and inserted into AscI linearized pKR92 (SEQ ID NO:27) to give ARALO79 (SEQ ID NO:71). The ARALO79 vector contains the following expression unit: BnSUS2-2 promoter::ZM-ODP1::phaseolin terminator.

Plasmid DNA of ARALO77, ARALO78, ARALO79 and ARALO80 were used for *Agrobacterium*-mediated transformation of *Arabidopsis* plants as described in Example 2.

#### Example 15

##### Analysis of Progeny Seed of *Arabidopsis* Plants Transformed with *Zea mays* ODP Under Control of Canola Sucrose Synthase Promoters

Oil content of progeny seed (e.g., T2 seed) of transgenic lines generated with ARALO78 and ARALO79 can be measured by NMR as described in Example 2. Progeny seed (e.g., T2 seed) of transgenic events generated with ARALO78 and ARALO79 are expected to show increased oil content when compared to seed of untransformed control plants grown alongside the transgenic events.

---

#### SEQUENCE LISTING

<160> NUMBER OF SEQ ID NOS: 74

<210> SEQ ID NO 1  
 <211> LENGTH: 5280  
 <212> TYPE: DNA  
 <213> ORGANISM: Artificial Sequence  
 <220> FEATURE:  
 <223> OTHER INFORMATION: vector pKS121/BS  
 <220> FEATURE:  
 <221> NAME/KEY: misc\_feature  
 <222> LOCATION: (4090)..(4090)  
 <223> OTHER INFORMATION: n is a, c, g, or t

<400> SEQUENCE: 1

```

ggccgccacc gcggtggagc tccagctttt gttcccttta gtgagggtta attgcgcgct    60
tggcgtaate atggtcatag ctgtttcctg tgtgaaattg ttatccgctc acaattccac    120
acaacatacg agccggaagc ataaagtgtg aagcctgggg tgctaatga gtgagctaac    180
tcacattaat tgcgttgccg tcaactgccc ctttccagtc gggaaacctg tegtgccagc    240
tgcattaatg aatcggccaa cgcgcgggga gagggcggtt gcgtattggg cgctcttccg    300
cttctcgcct cactgactcg ctgcgctcgg tegtctcgct gcggcgagcg gtatcagctc    360
actcaaaggc ggtaatacgg ttatccacag aatcagggga taacgcagga aagaacatgt    420
gagcaaaaag ccagcaaaaag gccaggaacc gtaaaaaggc cgcgttgctg gcgtttttcc    480
ataggctccg cccccctgac gagcatcaca aaaatcgacg ctcaagtcag aggtggcgaa    540
accgcagagg actataaaga taccaggcgt ttccccctgg aagctccctc gtgcgctctc    600
ctgttccgac cctgcgcgct accggatacc tgtccgcctt tctcccttcg ggaagcgtgg    660
cgctttctca tagctcaecg tgtaggatc tcagttcggt gtaggtcggt cgctccaagc    720
tgggctgtgt gcacgaaccc cccgttcagc ccgaccgctg cgcttatec ggtaactatc    780
gtcttgagtc caaccggta agacacgact tatcgccact ggcagcagcc actggtaaca    840
ggattagcag agcgaggtat gtaggcggtg ctacagagtt cttgaagtgg tggcctaact    900

```

-continued

---

acggctacac tagaaggaca gtatttggta tctgcgctct gctgaagcca gttaccttcg	960
gaaaaagagt tggtagctct tgatccggca aacaaaccac cgctggtagc ggtggttttt	1020
ttgtttgcaa gcagcagatt acgcgagaa aaaaaggatc tcaagaagat cctttgatct	1080
tttctacggg gtctgacgct cagtggaacg aaaactcacg ttaagggatt ttggtcata	1140
gattatcaaa aaggatcttc acctagatcc ttttaatta aaaatgaagt tttaatcaa	1200
tctaaagtat atatgagtaa acttggtctg acagttacca atgcttaatc agtgaggcac	1260
ctatctcagc gatctgtota tttcgttcat ccatagttgc ctgactcccc gtcgtgtaga	1320
taactacgat acgggagggc ttaccatctg gcccagtgcc tgcaatgata ccgagagacc	1380
cacgctcacc ggctccagat ttatcagcaa taaaccagcc agccggaagg gccgagcgca	1440
gaagtgggac tgcaacttta tccgctcca tccagttcat taattgttgc cgggaagcta	1500
gagtaagtag ttcgccagtt aatagtttgc gcaacgttgt tgccattgct acaggcatcg	1560
tggtgtcacg ctctctgttt ggtatggctt cattcagctc cggttcccaa cgatcaaggc	1620
gagttacatg atccccatg ttgtgcaaaa aagcggtagc ctcttcgggt cctccgatcg	1680
ttgtcagaag taagtgggac gcagtggtat cactcatggt tatggcagca ctgcataatt	1740
ctcttactgt catgccatcc gtaagatgct tttctgtgac tggtgagtag tcaaccaagt	1800
cattctgaga atagtgtatg cggcgaccga gttgctcttg cccggcgta atacgggata	1860
ataccgcgcc acatagcaga actttaaag tgctcatcat tggaaaacgt tcttcggggc	1920
gaaaactctc aaggatctta ccgctgttga gatccagttc gatgtaacct actcgtgcac	1980
ccaactgac ttcagcatct tttactttca ccagcgttcc tgggtgagca aaaacaggaa	2040
ggcaaaatgc cgcaaaaaag ggaataaggc gcacacggaa atgttgaata ctcatactct	2100
tcctttttca atattattga agcatttacc agggttattg tctcatgagc ggatacatat	2160
ttgaatgtat ttgaaaaat aaacaatag gggttccgag cacatttccc cgaaaagtgc	2220
cacctgacgc gccctgtagc ggcgcattaa gcgcggcggg tgtgggtggt acgagcagcg	2280
tgaccgctac acttgccagc gccctagcgc ccgctccttt cgctttcttc ccttcctttc	2340
tcgccagctt cgcggctttt ccccgtaag ctctaaatcg ggggctccct ttagggttcc	2400
gatttagtgc tttacggcac ctgcaccca aaaaacttga ttaggtgat ggttcacgta	2460
gtgggccatc gccctgtagc acggtttttc gccctttgac gttggagtcc acgttcttta	2520
atagtggact cttgttccaa actggaacaa cactcaacct tatctcggtc tattcttttg	2580
atttataagg gattttgccc atttcggcct attggttaaa aaatgagctg atttaacaaa	2640
aatttaacgc gaattttaac aaaatattaa cgcttacaat ttccattcgc cttcaggtc	2700
gcgcaactgt tgggaagggc gatcgggtgc ggcctcttcg ctattacgcc agctggcgaa	2760
agggggatgt gctgcaaggc gattaagttg ggtaacgcca gggttttccc agtcacgacg	2820
ttgtaaaacg acggccagtg agcgcgctga atacgactca ctatagggcg aattgggtac	2880
cgggcccccc ctgaggtgag acgggtatcga taagcttgat atcgaattcc tgcagcccg	2940
gggatccact agttctagag cggcgcgccc tcgacggata taatgagccc taaacaaga	3000
tgattaagta gtaattaata cgtactagta aaagtggcaa aagataacga gaaagaacca	3060
atttctttgc attcggcctt agcggaaaggc atataaagc tttgattatt ttatttagtg	3120
taatgatttc gtacaaccaa agcatttatt tagtactctc acacttgtgt cgcggccgct	3180
tgggggggta tggaaagact tcttagttag ttgtgtgaat aagcaatggt gggagaatcg	3240

-continued

---

```

ggactactta taggatagga ataaaacaga aaagtattaa gtgctaataa aatatttaga 3300
ctgataatta aaatcttcac gtatgtocac ttgatataaa aacgtcagga ataaaggaag 3360
tacagtagaa tttaaaggta ctctttttat atatacccggt gttctctttt tggctagcta 3420
gttgcataaa aaataatcta tattttttatc attatttttaa atatcttatg agatggtaaa 3480
tatttatcat aatttttttt actattattt attattttgtg tgtgtaatac atataagaagt 3540
taattacaaa ttttatttac tttttcatta ttttgatagc attcaccatt aattttagtg 3600
tattatttat aatagttcat tttaatcttt ttgtatatac tatgctgca gtactttttt 3660
cctacatata actactatta catttttatt atataatatt tttattaatg aattttcgtg 3720
ataatagta atattgttca ttattatttc agatttttta aaaatatttg tgttattatt 3780
tatgaaatat gtaatttttt tagtatttga ttttatgatg ataaagtgtt ctaaattcaa 3840
aagaaggggg aaagcgtaaa cattaaaaaa cgtcatcaaa caaaaacaaa atcttgttaa 3900
taaagataaa actgtttggc ttgatcactg ttatttcgta atataaaaaa attatttata 3960
tttatattgt tgacaaccaa atttgcctat caaatctaac caatataatg catgctggc 4020
aggtaaatgta ctaccatgaa cttaagtcac gacataataa accgtgaatc tgaccaatgc 4080
atgtacctan ctaaatgta tttgtgacac gaagcaaatg attcaattca caatggagat 4140
gggaaacaaa taatgaagaa cccagaacta agaaagcttt tctgaaaaat aaaataaagg 4200
caatgtcaaa agtatactgc atcatcagtc cagaaagcac atgatatttt tttatcagta 4260
tcaatgcagc tagttttatt ttacaatac gatatagcta gtttaaatat attgcagcta 4320
gatttataaa tatttgggtt attatttacc atttggtaa tcctgttttt agtattttag 4380
tttatatatg atgataatgt attcctcaatt taaaagaagg gaaataaatt taaacaagaa 4440
aaaaagtcac caaacaacaaa acaaatgaaa ggggtggaag atgttaccat gtaatgtgaa 4500
tgttacagta tttcttttat tatagagtta acaaatcaac taatagatt ttgttaataa 4560
tgataaaaata ttttttttat tattatttca taatataaaa atagtttact taatataaaa 4620
aaaattctat cgttcacaac aaagttggcc acctaattta accatgcatg taccatgga 4680
ccatattagg taaccatcaa acctgatgaa gagataaaga gatgaagact taagtcataa 4740
cacaacacca taaaaacaaa aaatacaatc aaccgtcaat ctgaccaatg catgaaaaag 4800
ctgcaatagt gagtggcgac acaaaagcaca tgattttctt acaacggaga taaaacaaa 4860
aaaatatttc atgaacaacc tagaacaat aaagctttta tataataaat atataataa 4920
ataaaggcta tggaataata tacttcaata tatttggatt aaataaattg ttggcggggt 4980
tgatataatt atacacacct aaagtcactt caatctcatt ttcactaac ttttattttt 5040
ttttctttt tatttatcat aaagagaata ttgataatat actttttaac atattttat 5100
gacatttttt attggtgaaa acttattaaa aatcataaat tttgtaagtt agatttattt 5160
aaagagttcc tcttcttatt ttaattttt taataaattt ttaataaact aaaatttgtg 5220
ttaaaatgt taaaaagtg tgttattaac ccttctctc gaggatccaa gcttggcgcg 5280

```

&lt;210&gt; SEQ ID NO 2

&lt;211&gt; LENGTH: 5968

&lt;212&gt; TYPE: DNA

&lt;213&gt; ORGANISM: Artificial Sequence

&lt;220&gt; FEATURE:

&lt;223&gt; OTHER INFORMATION: vector pDsRedxKS121/BS;DNA

&lt;220&gt; FEATURE:

&lt;221&gt; NAME/KEY: misc\_feature

&lt;222&gt; LOCATION: (5052)..(5052)

&lt;223&gt; OTHER INFORMATION: n is a, c, g, or t

-continued

&lt;400&gt; SEQUENCE: 2

catggcctcc tccgaggacg tcatcaagga gttcatgcgc ttcaaggtgc gcatggaggg 60  
 ctccgtgaac ggccacgagt tcgagatcga gggcgagggc gagggccgcc cctacgaggg 120  
 caccagacc gccaaagtga aggtgaccaa gggcgccccc ctgcccttcg cctgggacat 180  
 cctgtccccc cagttccagt acggctccaa ggtgtacgtg aagcaccgcc ccgacatccc 240  
 cgactacaag aagctgtcct tccccgaggg cttcaagtgg gagcgcgtga tgaacttcga 300  
 ggacggcggc gtggtgaccg tgaccagga ctctccctg caggacggct ccttcatcta 360  
 caagtggaag ttcacggcg tgaactccc ctccgacggc cccgtaatgc agaagaagac 420  
 tatgggctgg gaggcctcca ccgagcgcct gtacccccgc gacggcgtgc tgaagggcga 480  
 gatccacaag gccctgaagc tgaaggacgg cggccactac ctggtggagt tcaagtccat 540  
 ctacatggcc aagaagcccg tgcagctgcc cggctactac tacgtggact ccaagctgga 600  
 catcacctcc cacaacgagg actacaccat cgtggagcag tacgagcgcg ccgagggccg 660  
 ccaccacctg ttctgtagc ggcgggccgc gacacaagtg tgagagtact aaataaatgc 720  
 tttggttgta cgaatcatt aactaaata aaataatcaa agcttatata tgccttcgc 780  
 taaggccgaa tgcaagaaa ttggttcttt ctcgttatct tttgccctt ttactagtac 840  
 gtattaatta ctacttaate atctttgttt acggctcatt atatccgtcg acggcgcgcc 900  
 gctctagaac tagtggatcc cccgggctgc aggaattcga tatcaagctt atcgataccg 960  
 tcgacctcga gggggggccc ggtacceaat tcgccctata gtgagtcgta ttacgcgcgc 1020  
 tcaactggcg tcgttttaca acgtcgtgac tgggaaaacc ctggcgttac ccaacttaat 1080  
 cgccttgacg cacatcccc ttcgccagc tggcgttaata gcgaagaggg ccgcaccgat 1140  
 cgcccttccc aacagttgcg cagcctgaat ggcgaatgga aattgtaagc gttaaatatt 1200  
 tgtaaaaatt cgcgttaaat ttttgtaaa tcagctcatt ttttaacca taggcgaaa 1260  
 tcggcaaaat cccttataaa tcaaaagaat agaccgagat aggggtgagt gttgttcag 1320  
 tttggaacaa gagtccacta ttaagaacg tggactccaa cgtcaaaggg cgaaaaaccg 1380  
 tctatcaggg cgatggccca ctacgtgaac catcacccca atcaagtttt ttggggtcga 1440  
 ggtgccgtaa agcactaaat cgaacccta aaggagccc ccgatttaga gcttgacggg 1500  
 gaaagccggc gaacgtggcg agaaggaag ggaagaaagc gaaaggagcg ggcgctaggg 1560  
 cgctggcaag tgtagcggtc acgctgcgcg taaccaccac acccgccgcg cttaatgcgc 1620  
 cgctacaggg cgcgtcaggt ggcacttttc ggggaaatgt gcgcggaacc cctatttgtt 1680  
 tatttttcta aatacattca aatgatgatc cgctcatgag acaataaccc tgataaatgc 1740  
 ttcaataata ttgaaaaagg aagagtatga gtattcaaca tttccgtgc gcccttatc 1800  
 cctttttgc ggcattttgc ctctctgttt ttgctcacc agaaacgctg gtgaaagtaa 1860  
 aagatgctga agatcagttg ggtgcacgag tgggttacat cgaactggat ctcaacagcg 1920  
 gtaagatcct tgagagtttt cgcgccgaag aacgttttcc aatgatgagc acttttaag 1980  
 ttctgctatg tggcgcggta ttatccgta ttgacgcggc gcaagagcaa ctcggtcggc 2040  
 gcatacacta ttctcagaat gacttgggtg agtactcacc agtcacagaa aagcatctta 2100  
 cggatggcat gacagtaaga gaattatgca gtgctgccat aacctgagt gataaactg 2160  
 cggccaactt acttctgaca acgatcggag gaccgaagga gctaaccgct tttttgcaca 2220  
 acatggggga tcatgtaact cgccttgatc gttgggaacc ggagctgaat gaagccatac 2280

-continued

---

caaacgacga gcgtagacacc acgatgacct tagcaatggc aacaacgttg cgcaaactat	2340
taactggcga actacttact cttagcttccc ggcaacaatt aatagactgg atggaggcgg	2400
ataaagttgc aggaccactt ctgcgctcgg cccttccggc tggctggttt attgctgata	2460
aatctggagc cggtagagcgt gggctctcgc gtatcattgc agcactgggg ccagatggta	2520
agccctcccg tatcgtagtt atctacacga cggggagtca ggcaactatg gatgaacgaa	2580
atagacagat cgctgagata ggtgcctcac tgattaagca ttggtaactg tcagaccaag	2640
tttactcata tatactttag attgatttaa aacttcattt ttaatttaa aggatctagg	2700
tgaagatcct ttttgataat ctcatgacca aaatcccta acgtgagttt tcgttccact	2760
gagcgtcaga ccccgtagaa aagatcaaag gatcttcttg agatcctttt tttctgcgcg	2820
taatctgctg cttgcaaaaca aaaaaaccac cgctaccagc ggtggtttgt ttgccgatc	2880
aagagctacc aactcttttt ccgaaggtaa ctggcttcag cagagcgcag ataccaaata	2940
ctgtcctctt agttagcgcg tagttaggcc accacttcaa gaactctgta gcaccgcta	3000
catacctcgc tctgctaate ctgttaccag tggctgctgc cagtggcgat aagtcgtgtc	3060
ttaccggggt ggactcaaga cgatagttac cggataaagg gcagcggctc ggctgaacgg	3120
ggggttcctg cacacagccc agcttggagc gaacgaccta caccgaactg agatacctac	3180
agcgtgagct atgagaaagc gccacgcttc ccgaaggag aaaggcggac aggtatccgg	3240
taagcggcag ggtcggaaca ggagagcgcga cgaggagct tccaggggga aacgcctggt	3300
atctttatag tcctgtcggg tttcgcacc tctgacttga gcgtcgattt ttgtgatget	3360
cgtagggggg gcggagccta tggaaaaacg ccagcaacgc ggccttttta cggttcctgg	3420
ccttttctg gccttttctt cacatgttct ttcctgcgtt atcccctgat tctgtggata	3480
accgtattac cgcttttag tgagctgata ccgctcgcg cagccgaacg accgagcga	3540
gcgagtcagt gagcgaggaa gcggaagagc gcccaatacg caaacccctt ctccccgcgc	3600
gttgccgat tcattaatgc agctggcaag acaggtttcc cgactggaaa gcgggcagt	3660
agcgaacgc aattaatgt agttagctca ctcataggc accccaggct ttacacttta	3720
tgcttccggc tcgtatgttg tgtggaattg tgagcggata acaatttcac acaggaaca	3780
gctatgacca tgattacgcc aagcgcgcaa ttaaccctca ctaaagggaa caaaagctgg	3840
agctccaccg cgggtggcgc ccgcgccaag ctggatcct cgaagagaag ggttaataac	3900
acactttttt aacattttta acacaaattt tagttattta aaaatttatt aaaaaattta	3960
aaataagaag aggaactcct taaataaatc taacttaca aatttatgat ttttaataag	4020
ttttcaccaa taaaaatgt cataaaaaata tgttaaaaag tatattatca atattctctt	4080
tatgataaat aaaaagaaaa aaaaaataaa agttaagtga aatgagatt gaagtgactt	4140
taggtgtgta taaatatatc aaccccgcga acaatttatt taatccaaat atattgaagt	4200
atattattcc atagccttta tttatttata tatttattat ataaaagctt tatttgttct	4260
aggttgttca tgaatatatt ttttggtttt atctcogttg taagaaaatc atgtgctttg	4320
tgtcgcact cactattgca gctttttcat gcattggtea gattgacggt tgattgtatt	4380
tttgtttttt atggttttgt gttatgactt aagtcttcat ctctttatct ctccatcagg	4440
tttgatggtt acctaatatg gtccatgggt acatgcatgg ttaaattagg tggccaactt	4500
tgttgtaac gatagaatth tttttatatt aagtaacta tttttatatt atgaaataat	4560
aataaaaaaa atattttatc attattaaca aaatcatatt agttaatttg ttaactctat	4620
aataaaagaa atactgtaac atccacatta catggtaaca tctttccacc ctttcatttg	4680

-continued

---

```

ttttttgttt gatgactttt tttcttggtt aaatttattt cccttctttt aaatttgtaa 4740
tacattatca tcatatataa actaaaatac taaaaacagg attacacaaa tgataaataa 4800
taacacaaat atttataaat ctagctgcaa tatattttaa ctagctatat cgatattgta 4860
aaataaaaact agctgcattg atactgataa aaaaatatca tgtgctttct ggactgatga 4920
tgcagtatac ttttgacatt gcctttattt ttttttcag aaaagctttc ttagttctgg 4980
gttcttcatt atttgtttcc catctccatt gtgaattgaa tcatttgctt cgtgtcacia 5040
atacaattta gntaggtaca tgcattggtc agattcacgg tttattatgt catgacttaa 5100
gttcatggta gtacattacc tgccacgcat gcattatatt ggtagattt gataggcaaa 5160
tttggttgtc aacaatataa atataataa tgtttttata ttacgaaata acagtgatca 5220
aaacaaacag ttttatcttt attaacaaga ttttgttttt gtttgatgac gttttttaat 5280
gtttacgctt tcccccttct tttgaattta gaacacttta tcatcataaa atcaaatact 5340
aaaaaaaaata catatttcat aaataatac acaaatattt taaaaaatc tgaataata 5400
atgaacaata ttacatatta tcacgaaaat tcattaataa aaatattata taaataaaat 5460
gtaatagtag ttatatgtag gaaaaaagta ctgcaacgcat aatatatata aaaagattaa 5520
aatgaactat tataaataat aacactaaat taatggtgaa tcatatcaaa ataataaaaa 5580
agtaaaaaa atttgtaatt aacttctata tgtattacac acacaaataa taaataatag 5640
taaaaaaaat tatgataaat atttaccatc tcataagata tttaaaataa tgataaaaat 5700
atagattatt ttttatgcaa ctagctagcc aaaaagagaa cacgggtata tataaaaaga 5760
gtacctttaa attctactgt acttcttcta ttcctgacgt ttttatatca agtggacata 5820
cgtgaagatt ttaattatca gtctaaatat ttcattagca cttaatactt ttctgtttta 5880
ttctatctct ataagtagtc ccgattctcc caacattgct tattcacaca actaactaag 5940
aaagtcttcc atagccccc aagcggcc 5968

```

```

<210> SEQ ID NO 3
<211> LENGTH: 10058
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: vector pKS332
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: (1178)..(1178)
<223> OTHER INFORMATION: n is a, c, g, or t

```

```

<400> SEQUENCE: 3

```

```

gatcctcgaa gagaaggggt aataacacac tttttaaca ttttaacac aaattttagt 60
tatttaaaaa tttattaaaa aatttaaaat aagaagagga actctttaaa taaatctaac 120
ttacaaaatt tatgattttt aataagtttt caccaataaa aaatgtcata aaaatagtgt 180
aaaaagtata ttatcaatat tctctttatg ataaataaaa agaaaaaaaa aataaaagt 240
aagtgaaaaat gagattgaag tgactttagg tgtgtataaa tatatcaacc cgcaccaaa 300
tttatttaat ccaaatatat tgaagtatat tattccatag cctttattta tttatatatt 360
tattatataa aagctttatt tgttctaggt tgttcatgaa atattttttt ggttttatct 420
cggttgtaag aaaatcatgt gctttgtgtc gccactcact attgcagctt tttcatgcat 480
tggtcagatt gacggttgat tgtatttttg ttttttatgg ttttgtgtta tgacttaagt 540
cttcatctct ttatctcttc atcaggtttg atggttaect aatatggctc atgggtacat 600

```

-continued

---

gcatggttaa	attaggtggc	caactttgtt	gtgaacgata	gaattttttt	tatattaagt	660
aaactatttt	tatattatga	aataataata	aaaaaaat	tttatcatta	ttaacaaaat	720
catattagtt	aatttggtta	ctctataata	aaagaaatc	tgtaacattc	acattacatg	780
gtaacatctt	tccacccttt	catttgttt	ttgttgatg	actttttttc	ttgtttaaat	840
ttatttcctt	tcttttaaat	ttggaataca	ttatcatcat	atataaacta	aaatactaaa	900
aacaggatta	cacaaatgat	aaataataac	acaaatattt	ataaatctag	ctgcaatata	960
tttaaactag	ctatatcgat	attgtaaaat	aaaactagct	gcattgatac	tgataaaaaa	1020
atatcatgtg	ctttctggac	tgatgatgca	gtatactttt	gacattgcct	ttattttatt	1080
tttcagaaaa	gctttcttag	ttctgggttc	ttcattattt	gtttcccatc	tccattgtga	1140
attgaatcat	ttgcttcgtg	tcacaaatac	aatttagnta	ggtacatgca	ttggtcagat	1200
tcacggttta	ttatgtcatg	acttaagttc	atggtagtac	attacctgcc	acgcatgcat	1260
tatattgggt	agatttgata	ggcaaatttg	gttgcaaca	atataaatat	aaataatggt	1320
tttatattac	gaaataacag	tgatcaaaac	aaacagtttt	atctttatta	acaagatttt	1380
gtttttgttt	gatgacgttt	tttaatgttt	acgctttccc	ccttcttttg	aatttagaac	1440
actttatcat	cataaaatca	aatactaaaa	aaattacata	tttcataaat	aataacacaa	1500
atatttttaa	aaaatctgaa	ataataatga	acaatattac	atattatcac	gaaaattcat	1560
taataaaaaat	attatataaa	taaaatgtaa	tagtagttat	atgtaggaaa	aaagtactgc	1620
acgcataata	tatacaaaaa	gattaaaaatg	aactattata	aataataaca	ctaaattaat	1680
ggtgaatcat	atcaaaataa	tgaaaaagta	aataaaattt	gtaattaact	tctatatgta	1740
ttacacacac	aaataataaa	taatagtaaa	aaaaattatg	ataaatattt	accatctcat	1800
aagatattta	aaataatgat	aaaaatatag	attatttttt	atgcaactag	ctagcaaaaa	1860
agagaacacg	ggtatatata	aaaagagtac	ctttaaattc	tactgtactt	cctttattcc	1920
tgacgttttt	atatcaagtg	gacatactg	aagattttta	ttatcagttc	aaatatttca	1980
ttagcaacta	atacttttct	gttttattcc	tatcctataa	gtagtccga	ttctcccaac	2040
attgcttatt	cacacaacta	actaagaaag	tcttccatag	ccccccaagc	ggcccatggc	2100
ctcctccgag	gacgtcatca	aggagtcat	gcgcttcaag	gtgcgcatgg	agggtccgt	2160
gaacggccac	gagttcgaga	tcgagggcga	ggcgagggc	cgcccctacg	agggcaccca	2220
gaccgccaag	ctgaaggtga	ccaagggcgg	ccccctgccc	ttcgctggg	acatcctgtc	2280
ccccagttc	cagtacggct	ccaaggtgta	cgtgaagcac	cccgcgcaca	tccccgacta	2340
caagaagctg	tccttccccg	agggttcaa	gtgggagcgc	gtgatgaact	tcgaggacgg	2400
cggcgtgggt	accgtgacct	aggactcctc	cctgcaggac	ggctccttca	tctacaaggt	2460
gaagttcatc	ggcgtgaact	tccccctcga	cggccccgta	atgcagaaga	agactatggg	2520
ctgggaggcc	tccaccgagc	gctgtacctc	ccgcgacggc	gtgctgaagg	gcgagatcca	2580
caaggccctg	aagctgaagg	acggcggcca	ctacctggtg	gagttcaagt	ccatctacat	2640
ggccaagaag	cccgtgcagc	tgccccgcta	ctactactgt	gactccaagc	tgacatcac	2700
ctcccacaac	gaggactaca	ccatcgtgga	gcagtacgag	cgcgcgaggg	gccgccacca	2760
cctgttcctg	tageggccgg	cgcgcacaca	agtgtagag	tactaaataa	atgctttggt	2820
tgtaacgaaat	cattacaacta	aataaaataa	tcaaagctta	tatatgcctt	ccgctaagge	2880
cgaatgcaaa	gaaattgggt	ctttctcgtt	atcttttgcc	acttttacta	gtacgtatta	2940
attactactt	aatcatcttt	gtttacggct	cattatatcc	gtcgacggcg	cgccgctcta	3000

-continued

---

gaactagtgg atccgtcgac ggcgcgccc atcatccgga tatagttcct cctttcagca	3060
aaaaaccct caagaccogt ttagaggccc caaggggta tgctagtat tgctcagcgg	3120
tggcagcagc caactcagct tcctttcggg ctttgtagc agccggatcg atccaagctg	3180
tacctcacta ttctttgccc ctccgacgag tgctggggcg tcggtttcca ctatcggcga	3240
gtacttttac acagccatcg gtccagacgg ccgcgcttct gcgggcgatt tgtgtacgcc	3300
cgacagtccc ggctccggat cggacgattg cgtcgcacatg accctgcgcc caagctgcat	3360
catcgaaatt gccgtcaacc aagctctgat agagttggtc aagaccaatg cggagcatat	3420
acgcccggag ccgcggcgat cctgcaagct ccggatgcct ccgctcgaag tagcgcgtct	3480
gctgtccat acaagccaac cacggcctcc agaagaagat gttggcgacc tcgtattggg	3540
aatccccgaa catcgcctcg ctccagtcaa tgaccgctgt tatgcggccca ttgtccgtca	3600
ggacattgtt ggagccgaaa tccgcgtgca cgaggtgccc gacttcgggg cagtccctcg	3660
cccaaagcat cagctcatcg agagcctcgc cgacggacgc actgacggtg tcgtccatca	3720
cagtttgcca gtgatacaca tggggatcag caatcgcgca tatgaaatca cgccatgtag	3780
tgtattgacc gattccttgc ggtccgaatg ggccgaacct gctcgtctgg ctaagatcgg	3840
ccgcagcgat cgcattcata gcctccgca ccggctgcag aacagcgggc agttcggttt	3900
caggcaggtc ttgcaacgtg acaccctgtg cacggcggga gatgcaatag gtcaggctct	3960
cgctgaattc cccaatgtca agcacttccg gaatcgggag ccgcggccgat gcaaagtgcc	4020
gataaacata acgatctttg tagaaacct ccggcgcagct atttaccgc aggacatatc	4080
cacgcccctc tacatcgaag ctgaaagcac gagattcttc gccctccgag agctgcatca	4140
ggtcggagac gctgtcgaac ttttcgatca gaaacttctc gacagacgtc gcggtgagtt	4200
caggcttttc catgggtata tctccttctt aaagttaaac aaaattattt ctagagggaa	4260
accgttgtgg tctccctata gtgagtcgta ttaatttcgc gggatcgaga tcgatccaat	4320
tccaatcca caaaaatctg agcttaacag cacagttgct cctctcagag cagaatcggg	4380
tattcaacac cctcatatca actactacgt tgtgtataac ggtccacatg ccggtatata	4440
cgatgactgg ggtgtgatac aggcggcaac aaacggcggt cccggagttg cacacaagaa	4500
atgtgccact attacagagg caagagcagc agctgacgcg tacacaacaa gtcagcaaac	4560
agacaggttg aacttcaccc ccaaaaggaga agctcaactc aagcccaaga gctttgctaa	4620
ggcctaaca agcccaccaa agcaaaaagc ccactggctc acgctaggaa ccaaaaggcc	4680
cagcagtgat ccagcccaa aagagatctc ctttgccccg gagattacaa tggacgattt	4740
cctctatctt tacgatctag gaaggaagt cgaagggtgaa ggtgacgaca ctatgttcac	4800
cactgataat gagaagggtta gcctcttcaa tttcagaaag aatgctgacc cacagatggt	4860
tagagaggcc tacgcagcag gtctcatcaa gacgatctac ccgagtaaca atctccagga	4920
gatcaaatac cttccaaga aggttaaaga tgcagtaaaa agattcagga ctaattgcat	4980
caagaacaca gagaagaca tattttctaa gatcagaagt actattccag tatggacgat	5040
tcaaggcttg cttcataaac caaggcaagt aatagagatt ggagtctcta aaaaggtagt	5100
tcctactgaa tctaaggcca tgcattggagt ctaagattca aatcaggat ctaacagaac	5160
tcgccgtgaa gactggcgaa cagttcatac agagtctttt acgactcaat gacaagaaga	5220
aaatctctgt caacatggtg gagcagcaca ctctggtcta ctccaaaaat gtcaaaagata	5280
cagtctcaga agacaaaagg gctattgaga cttttcaaca aaggataatt tcgggaaacc	5340

-continued

---

tctctcgatt	ccattgcccc	gctatctgtc	acttcatcga	aaggacagta	gaaaaggaag	5400
gtggctcccta	caaatgccat	cattgcgata	aaggaaagc	tatcattcaa	gatgcctctg	5460
cgcacagtgg	tcccaaagat	ggacccccac	ccacgaggag	catcgtggaa	aaagaagacg	5520
ttccaaccac	gtcttcaaag	caagtggatt	gatgtgacat	ctccactgac	gtaagggatg	5580
acgcacaatc	ccactatcct	tcgcaagacc	cttcctctat	ataaggaagt	tcatttcatt	5640
tggagaggac	acgctcgagc	tcattttctc	attacttcag	ccataacaaa	agaactcttt	5700
tctcttctta	ttaaaccatg	aaaaagcctg	aactcaccgc	gacgtctgtc	gagaagtttc	5760
tgatcgaaaa	gttcgacagc	gtctccgacc	tgatgcagct	ctcggagggc	gaagaatctc	5820
gtgctttcag	cttcgatgta	ggagggcgtg	gatatgtcct	gcgggtaaat	agctgcgccc	5880
atggtttcta	caaagatcgt	tatgtttatc	ggcactttgc	atcggccgcg	ctcccgatcc	5940
cggaagtgtc	tgacattggg	gaattcagcg	agagcctgac	ctattgcatc	tcccgcctg	6000
cacaggggtg	cacgttgcaa	gacctgcctg	aaaccgaact	gcccgcctgt	ctgcagccgg	6060
tcgaggaggc	catggatgcg	atcgtcgcgg	ccgatcttag	ccagacgagc	gggttcggcc	6120
cattcggacc	gcaaggaatc	ggtcaataca	ctacatggcg	tgatttcata	tgcgcgattg	6180
ctgatcccca	tgtgtatcac	tggcaaacctg	tgatggacga	caccgtcagt	gcgtccgtcg	6240
cgcaggctct	cgatgagctg	atgctttggg	ccgaggactg	ccccgaagtc	cggcacctcg	6300
tgcacgcgga	tttcggctcc	aacaatgtcc	tgacggacaa	tggccgcata	acagcggta	6360
ttgactggag	cgaggcgatg	ttcgggggatt	cccaatacga	ggtcgccaac	atcttctctc	6420
ggaggccctg	gttggcttgt	atggagcagc	agacgcgcta	cttcgagcgg	aggcatccgg	6480
agcttgacag	atcgcgcgcg	ctccggcgtc	atatgctccg	cattggtctt	gaccaactct	6540
atcagagctt	ggttgacggc	aatttcgatg	atgcagcttg	ggcgcagggg	cgatgcgacg	6600
caatcgtccg	atccggagcc	gggactgtcg	ggcgtacaca	aatcgcgccg	agaagcgcgg	6660
ccgtctggac	cgatggctgt	gtagaagtac	tcgcccgatg	tggaaaaccg	cgccccagca	6720
ctcgtccgag	ggcaaaggaa	tagtgaggta	cctaagaag	gagtgcgtcg	aagcagatcg	6780
ttcaaacatt	tggcaataaa	gtttcttaag	attgaatcct	gttgccggtc	ttgcgatgat	6840
tatcatataa	tttctgttga	attacgttaa	gcatgtaata	attaacatgt	aatgcatgac	6900
gttatttatg	agatgggttt	ttatgattag	agtcccgcaa	ttatacattt	aatcgcgat	6960
agaaaacaaa	atatagcgcg	caaacctagga	taaattatcg	cgcgcgggtg	catctatggt	7020
actagatcga	tgtcgaatct	gatcaacctg	cattaatgaa	tcggccaacg	cgcggggaga	7080
ggcggtttgc	gtattggggc	ctcttcgctc	tctctcgtca	ctgactcgtc	gcgctcggtc	7140
gttcggctgc	ggcgcggcgt	atcagctcac	tcaaaggcgg	taatacgggt	atccacagaa	7200
tcaggggata	acgcaggaaa	gaacatgtga	gcaaaagccc	agcaaaaggc	caggaaccgt	7260
aaaaaggccg	cgttgctggc	gtttttccat	aggctccgcc	cccctgacga	gcatcacaaa	7320
aatcgcgctc	caagtcagag	gtggcgaaac	ccgacaggac	tataaagata	ccaggcgttt	7380
ccccctggaa	gctccctcgt	gcgctctcct	gttccgaccc	tgccgcttac	cggatacctg	7440
tccgcctttc	tcccttcggg	aagcgtggcg	ctttctcaat	gctcacgctg	taggtatctc	7500
agttcgggtg	aggctcgttc	ctccaagctg	ggctgtgtgc	acgaaccccc	cgttcagccc	7560
gaccgctcgc	ccttatccgg	taactatcgt	cttgagtcca	acccggtaag	acacgactta	7620
tcgccactgg	cagcagccac	tggtaacagg	attagcagag	cgaggtatgt	aggcgggtct	7680
acagagttct	tgaagtgggtg	gcctaactac	ggctacacta	gaaggacagt	atttggatc	7740

-continued

---

tgcgctctgc	tgaagccagt	taccttcgga	aaaagagttg	gtagctcttg	atccggcaaa	7800
caaaccacog	ctggtagcgg	tggttttttt	gtttgcaagc	agcagattac	gcgcagaaaa	7860
aaaggatctc	aagaagatcc	tttgatcttt	tctacggggg	ctgacgctca	gtggaacgaa	7920
aactcacggt	aagggatttt	ggtcatgaca	ttaacctata	aaaataggcg	tatcacgagg	7980
ccctttcgtc	tcgcgcgttt	cggtgatgac	ggtgaaaacc	tctgacacat	gcagctcccc	8040
gagacgggtc	cagcttgctc	gtaagcggat	gccggggagc	gacaagcccc	tcagggcgcg	8100
tcagcgggtg	ttggcgggtg	tcggggctgg	cttaactatg	cggcatcaga	gcagattgta	8160
ctgagagtgc	accatattga	catattgtcg	ttagaacgcg	gctacaatta	atacataacc	8220
ttatgtatca	tacacatacg	atttaggtga	cactatagaa	cggcgcgcca	agcttttgat	8280
ccatgccctt	catttgccgc	ttattaatta	at ttggtaac	agtcgctact	aatcagttac	8340
ttatccttcc	cccatcataa	ttaatcttgg	tagtctcgaa	tgccacaaca	ctgactagtc	8400
tcttggatca	taagaaaaag	ccaaggaaca	aaagaagaca	aaacacaatg	agagtatcct	8460
ttgcatagca	atgtctaagt	tcataaaatt	caaacaaaaa	cgcaatcaca	cacagtggac	8520
atcacttatc	cactagctga	tcaggatcgc	cgcgtcaaga	aaaaaaaaact	ggaccccaaa	8580
agccatgcac	aacaacacgt	actcacaag	gtgtcaatcg	agcagcccaa	aacattcacc	8640
aactcaaccc	atcatgagcc	ctcacatttg	ttgtttctaa	cccaacctca	aactcgtatt	8700
ctcttccgcc	acctcatttt	tgtttatttc	aacacccgtc	aaactgcatg	ccaccccgty	8760
gccaaatgtc	catgcatggt	aacaagacct	atgactataa	atagctgcaa	tctcggccca	8820
ggttttcatc	atcaagaacc	agttcaatat	cctagtcac	cgtattaaag	aatttaagat	8880
atactgcggc	cgcaagtatg	aactaaaatg	catgtaggty	taagagctca	tggagagcat	8940
ggaatattgt	atccgaccat	gtaacagtat	aataactgag	ctccatctca	cttcttctat	9000
gaataaacaa	aggatgttat	gatataattaa	cactctatct	atgcacctta	ttgttctatg	9060
ataaatttcc	tcttattatt	ataaatcctc	tgaatcgtga	cggcttatgg	aatgcttcaa	9120
atagtacaaa	aacaaatgty	tactataaga	ctttctaaac	aattctaacc	ttagcattgt	9180
gaacgagaca	taagtgttaa	gaagacataa	caattataat	ggaagaagtt	tgtctccatt	9240
tatatattat	atattaccca	cttatgtatt	atattaggat	gttaaggaga	cataacaatt	9300
ataaagagag	aagtttgat	ccatttatat	attatatact	accatttat	atattatact	9360
tatccactta	tttaatgtct	ttataagggt	tgatccatga	tattttctaat	at tttagttg	9420
atatgtatat	gaaagggtag	tatttgaact	ctcttactct	gtataaagg	tggatcatcc	9480
ttaaagtggg	tctatttaat	tttattgctt	cttacagata	aaaaaaaaat	tatgagttgg	9540
tttgataaaa	tattgaagga	tttaaaataa	taataaataa	catataatat	atgtatataa	9600
at ttattata	atataacatt	tatctataaa	aaagtaataa	ttgtcataaa	tctatacaat	9660
cgtttagcct	tgctggacga	atctcaatta	tttaaacgag	agtaaacata	tttgactttt	9720
tggttattta	acaaattatt	atttaacact	atatgaaatt	ttttttttta	tcagcaaaga	9780
ataaaattaa	attaagaagg	acaatggtyt	cccaatcctt	atacaaccaa	cttcacaag	9840
aaagtcaagt	cagagacaac	aaaaaaaaaa	gcaaaggaaa	ttttttaatt	tgagttgtct	9900
tgtttgctgc	ataatttatg	cagtaaaaca	ctacacataa	cccttttagc	agtagagcaa	9960
tggttgacog	tgtgcttagc	ttcttttatt	ttattttttt	atcagcaaag	aataaataaa	10020
ataaaatgag	acacttcagg	gatgtttcaa	caagctty			10058

-continued

---

<210> SEQ ID NO 4  
 <211> LENGTH: 34  
 <212> TYPE: DNA  
 <213> ORGANISM: Artificial Sequence  
 <220> FEATURE:  
 <223> OTHER INFORMATION: PCR primer MWG345  
  
 <400> SEQUENCE: 4  
  
 gaattcgcgg cgcgatggag agatctcaac ggca 34

<210> SEQ ID NO 5  
 <211> LENGTH: 34  
 <212> TYPE: DNA  
 <213> ORGANISM: Artificial Sequence  
 <220> FEATURE:  
 <223> OTHER INFORMATION: PCR primer MWG346  
  
 <400> SEQUENCE: 5  
  
 gaattcgcgg cgcgcttagtt gcacacactg atca 34

<210> SEQ ID NO 6  
 <211> LENGTH: 11251  
 <212> TYPE: DNA  
 <213> ORGANISM: Artificial Sequence  
 <220> FEATURE:  
 <223> OTHER INFORMATION: vector pKS336  
 <220> FEATURE:  
 <221> NAME/KEY: misc\_feature  
 <222> LOCATION: (3542)..(3542)  
 <223> OTHER INFORMATION: n is a, c, g, or t  
  
 <400> SEQUENCE: 6  
  
 ggccgcatgg agagatctca acggcagtct cctccgccac cgtcgcctgc ctctctctcg 60  
 tcctccgtct cgcgggacac cgtcctcgtc cctcccgaa agaggcggag ggcggcgacg 120  
 gccaaaggcgc ggcggagacc taataagagg atccgcaagg accccgccgc cgccgcccgc 180  
 gggaagagga gctccgtcta caggggagtc accaggcaca ggtggaacgg caggttcgag 240  
 gcgcatctct gggacaagca ctgcctcgtc gcgctccaca acaagaagaa aggcaggcaa 300  
 gtctacctgg gggcgatga cagcgaggag gcagctgctc gtgcctatga cctcgcagct 360  
 ctcaagtact ggggtcctga gactctgctc aactccctg tggaggatta ctccagcgag 420  
 atgccggaga tggaggccgt gtcccgggag gactacctgg cctccctccg ccgagggagc 480  
 agcggctctc ccagggcgt ctccaagta agaggcgtc ccaggcatca ccacaacggg 540  
 aggtgggagg cacggattgg gcgagctctt gggaacaagt acctctactt gggaacattt 600  
 gacactcaag aagaggcagc caaggcctat gaccttgccg ccattgaata ccgtggcgtc 660  
 aatgctgtaa ccaacttoga catcagctgc tacctggacc acccgctgtt cctggcacag 720  
 ctccaacagg agccacaggt ggtgccggca ctcaaccaag aacctcaacc tgatcagagc 780  
 gaaaccggaa ctacagagca agagccggag tcaagcgaag ccaagacacc ggatggcagt 840  
 gcagaaccgc atgagaacgc ggtgcctgac gacaccgcgg agcccctcac cacagtcgac 900  
 gacagcatcg aagaggcctt gtaggaccct tgcattgatt acgagctaga caccatgtcg 960  
 agaccaaact ttggcagctc aatcaatctg agcgagtggg tcgctgacgc agacttcgac 1020  
 tgcaacatcg gatgcctgtt cgtgggtgt tctgcccgtg acgaaggaag caaggatggt 1080  
 gtaggtctgg cagatttcag tctggttgag gcaggtgatg tccagctgaa ggatgttctt 1140  
 tcggatatgg aagaggggat acaacctcca gcgatgatca gtgtgtgcaa cgcggccgca 1200  
 agtatgaact aaaatgcatg taggtgtaag agctcatgga gagcatggaa tattgtatcc 1260

-continued

---

gaccatgtaa	cagtataata	actgagctcc	atctcacttc	ttctatgaat	aaacaaagga	1320
tgttatgata	tattaacact	ctatctatgc	accttattgt	tctatgataa	atttcctctt	1380
attattataa	atcatctgaa	tcgtgacggc	ttatggaatg	cttcaaatag	tacaaaaaca	1440
aatgtgtact	ataagacttt	ctaaacaatt	ctaaccctag	cattgtgaa	gagacataag	1500
tgtaagaag	acataacaat	tataatggaa	gaagtttgtc	tccatttata	tattatata	1560
taccactta	tgtattatat	taggatgtta	aggagacata	acaattataa	agagagaagt	1620
ttgtatccat	ttatataatta	tatactacc	atttatata	tatacttate	cacttattta	1680
atgtctttat	aaggtttgat	ccatgatatt	tctaataatt	tagttgat	gtatatgaaa	1740
gggtactatt	tgaactctct	tactctgtat	aaaggttgg	tcatccttaa	agtggttcta	1800
tttaatttta	ttgcttctta	cagataaaaa	aaaaattatg	agttggtttg	ataaaatatt	1860
gaaggattta	aaataataat	aaataacata	taatatatgt	atataaattt	attataatat	1920
aacatttate	tataaaaaag	taaatattgt	cataaatcta	tacaatcgtt	tagccttgct	1980
ggacgaatct	caattattta	aacgagagta	aacatatttg	actttttggt	tatttaacaa	2040
attattattt	aacactatat	gaaatttttt	tttttatcag	caaagaataa	aattaaatta	2100
agaaggacaa	tggtgtccca	atccttatac	aaccaacttc	cacaagaaag	tcaagtcaga	2160
gacaacaaaa	aaacaagcaa	aggaaatttt	ttaatttgag	ttgtcttgtt	tgctgcataa	2220
tttatgcagt	aaaacactac	acataaccct	tttagcagta	gagcaatggt	tgaccgtgtg	2280
cttagcttct	tttattttat	ttttttatca	gcaaagaata	aataaaataa	aatgagacac	2340
ttcagggatg	tttcaacaag	cttggatcct	cgaagagaag	ggtaataaac	acactttttt	2400
aacattttta	acacaaattt	tagttattta	aaaatttatt	aaaaattta	aaataagaag	2460
aggaactcct	taaataaatc	taacttacia	aatttatgat	ttttaataag	ttttcaccaa	2520
taaaaaatgt	cataaaaaata	tgtaaaaaag	tatattatca	atattctctt	tatgataaat	2580
aaaaagaaaa	aaaaaataaa	agttaagtga	aatgagatt	gaagtgactt	taggtgtgta	2640
taaatatate	aaccccgcca	acaatttatt	taatccaat	atattgaagt	atattattcc	2700
atagccttta	tttatttata	tatttattat	ataaaagctt	tatttgttct	aggttgttca	2760
tgaaatattt	ttttggtttt	atctccgttg	taagaaaatc	atgtgctttg	tgtgcccact	2820
cactattgca	gctttttcoat	gcattggtea	gattgacggt	tgattgtatt	tttgtttttt	2880
atggttttgt	gttatgactt	aagtcttcat	ctctttatct	cttcatcagg	tttgatggtt	2940
acctaatatg	gtccatgggt	acatgcatgg	ttaaattagg	tggccaactt	tgttgtgaa	3000
gatagaattt	tttttatatt	aagtaaacata	tttttatatt	atgaaataat	aataaaaaaa	3060
atattttate	attattaaca	aaatcatatt	agttaatttg	ttaactctat	ataaaaagaa	3120
atactgtaac	attcacatta	catggtaaca	tctttccacc	ctttcatttg	tttttgttt	3180
gatgactttt	tttcttgttt	aaatttattt	cccttctttt	aaatttggaa	tacattatca	3240
tcatatataa	actaaaatac	taaaaacagg	attacacaaa	tgataaataa	taacacaaat	3300
atttataaat	ctagctgcaa	tatattttaa	ctagctatat	cgatattgta	aaataaaact	3360
agctgcattg	atactgataa	aaaaatatca	tgtgctttct	ggactgatga	tgcagtatac	3420
ttttgacatt	gcctttattt	tatttttcag	aaaagcttcc	ttagttctgg	gttcttcatt	3480
atgtgtttcc	catctccatt	gtgaattgaa	tcatttgctt	cgtgtcacia	atacaattta	3540
gntaggatca	tgcattggtc	agattcacgg	tttattatgt	catgacttaa	gttcatggta	3600

-continued

---

gtacattacc	tgccacgcat	gcattatatt	ggttagattt	gataggcaaa	tttggttgtc	3660
aacaatataa	atataaataa	tgtttttata	ttacgaaata	acagtgatca	aaacaaacag	3720
ttttatcttt	attaacaaga	ttttgtttt	gtttgatgac	gttttttaat	gtttacgctt	3780
tcccccttct	tttgaattta	gaacacttta	tcatcataaa	atcaaatact	aaaaaaatta	3840
catatttcat	aaataataac	acaatattt	ttaaaaaatc	tgaataata	atgaacaata	3900
ttacatatta	tcacgaaaat	tcattaataa	aaatattata	taaataaaat	gtaatagtag	3960
ttatatgtag	gaaaaagta	ctgcacgcat	aatatataca	aaaagattaa	aatgaactat	4020
tataaataat	aacactaat	taatggtgaa	tcatatcaaa	ataatgaaa	agtaaataaa	4080
atgtgtaatt	aacttctata	tgtattacac	acacaaataa	taaataatag	taaaaaaat	4140
tatgataaat	atttaccatc	tcataagata	tttaaaataa	tgataaaat	atagattatt	4200
ttttatgcaa	ctagctagcc	aaaaagagaa	cacgggtata	tataaaaaga	gtacctttaa	4260
atttctactgt	acttccttta	ttcctgacgt	ttttatatca	agtgagacata	cgtgaagatt	4320
ttaattatca	gtctaaatat	ttcattagca	cttaataactt	ttctgtttta	ttcctatcct	4380
ataagtagtc	ccgattctcc	caacattgct	tattcacaca	actaactaag	aaagtcttcc	4440
atagccccc	aagcggccca	tggcctctc	cgaggacgtc	atcaaggagt	tcatgcgctt	4500
caaggtgctc	atggagggtc	ccgtgaacgg	ccacgagttc	gagatcgagg	gcgagggcga	4560
gggcccgc	tacgagggca	cccagaccgc	caagctgaag	gtgaccaagg	gcgggcccct	4620
gcccctcgcc	tgggacatcc	tgtcccccca	gttccagtag	ggctccaagg	tgtacgtgaa	4680
gcaccccgcc	gacatccccg	actacaagaa	gctgtccttc	cccgagggtc	tcaagtggga	4740
gcgcgtgatg	aacttcgagg	acggcggcgt	ggtgaccgtg	accaggact	cctccctgca	4800
ggacggctcc	ttcatctaca	aggtgaagtt	catcggcgtg	aacttcccct	ccgacggccc	4860
cgtaatgcag	aagaagacta	tgggctggga	ggcctccacc	gagcgcctgt	acccccgcga	4920
cggcgtgctg	aagggcgaga	tccacaaggc	cctgaagctg	aaggacggcg	gccactacct	4980
ggtggagtcc	aagtccatct	acatggocaa	gaagcccgtg	cagctgcccg	gctactacta	5040
cgtggactcc	aagctggaca	tcacctcca	caacgaggac	tacaccatcg	tggagcagta	5100
cgagcgcgcc	gagggccgcc	accacctgtt	cctgtagcgg	ccggcccgcga	cacaagtgtg	5160
agagtactaa	ataaatgctt	tggttgtacg	aaatcattac	actaaataaa	ataatcaaag	5220
cttatatatg	ccttccgcta	aggcccgaatg	caaagaaatt	ggttctttct	cgttatcttt	5280
tgccactttt	actagtacgt	attaattact	acttaatcat	ctttgtttac	ggctcattat	5340
atccgtcgac	ggcgcgcgcg	tctagaacta	gtggatccgt	cgacggcgcg	cccgatcacc	5400
cggatatagt	tcctccttcc	agcaaaaaac	ccctcaagac	ccgtttagag	gccccaaagg	5460
gttatgctag	ttattgctca	gcgggtggcag	cagccaactc	agcttccttt	cgggctttgt	5520
tagcagccgg	atcgatccaa	gctgtaoctc	actattcctt	tgcctcggga	cgagtgtctgg	5580
ggcgtcggtt	tccactatcg	gcgagtactt	ctacacagcc	atcggctccag	acggccgcgc	5640
ttctgcgggc	gatttgtgta	cgccccacag	tcccggctcc	ggatcggacg	attgcgtcgc	5700
atcgaccctg	cgcccaagct	gcacatcga	aattgcccgc	aaccaagctc	tgatagagtt	5760
ggtcaagacc	aatgcccagc	ataacgccc	ggagccgcgg	cgatccctgca	agctccggat	5820
gcctccgctc	gaagtagcgc	gtctgtgct	ccatacaagc	caaccacggc	ctccagaaga	5880
agatgttggc	gacctcgat	tgggaatccc	cgaacatcgc	ctcgtccag	tcaatgaccg	5940
ctgttatgog	gccattgtcc	gtcaggacat	tgttggagcc	gaaatccgcg	tgcacgaggt	6000

-continued

---

gccggacttc	ggggcagtc	tcggcccaaa	gcacagctc	atcgagagcc	tgcgcgacgg	6060
acgcactgac	ggtgtcgtcc	atcacagttt	gccagtgata	cacatgggga	tcagcaatcg	6120
cgcatatgaa	atcacgccat	gtagtgatt	gaccgattcc	ttgcggtc	aatgggcccga	6180
acccgctcgt	ctggctaaga	tcggccgcag	cgatcgatc	catagcctcc	gcgaccggct	6240
gcagaacagc	gggcagttcg	gtttcaggca	ggtcttgc	cgtagacccc	tgtgcacggc	6300
gggagatgca	ataggtcagg	ctctcgtga	attccccaat	gtcaagcact	tccggaatcg	6360
ggagcgcggc	cgatgcaaa	tgccgataaa	cataacgatc	ttttagaaa	ccatcgccgc	6420
agctatttac	ccgcaggaca	tatccacgcc	ctcctacatc	gaagctgaaa	gcacgagatt	6480
cttcgccctc	cgagagctgc	atcaggtcgg	agacgctg	gaacttttcg	atcagaaact	6540
tctcgacaga	cgtcgcgggt	agttcaggct	tttccatggg	tatatctcct	tcttaaagtt	6600
aaacaaaatt	atttctagag	ggaacccgtt	gtggtctccc	tatagtgagt	cgtattaatt	6660
tcgcgggatc	gagatcgatc	caattccaat	cccacaaaa	tctgagctta	acagcacagt	6720
tgctcctctc	agagcagaat	cgggtattca	acaccctcat	atcaactact	acgttggtga	6780
taacgggtcca	catgccggta	tatacgatga	ctggggttgt	acaaaggcgg	caacaaacgg	6840
cgttcccgga	gttgcacaca	agaaatttgc	cactattaca	gaggcaagag	cagcagctga	6900
cgcgtagaca	acaagtcagc	aaacagacag	gttgaacttc	atccccaaag	gagaagctca	6960
actcaagccc	aagagctttg	ctaaggccct	aacaagccca	ccaaagcaaa	aagcccactg	7020
gctcacgcta	ggaacaaaa	ggcccagcag	tgatccagcc	ccaaaagaga	tctcctttgc	7080
cccggagatt	acaatggagc	atttctctca	tctttacgat	ctaggaagga	agttcgaagg	7140
tgaaggtgac	gacactatgt	tcaccactga	taatgagaag	gttagcctct	tcaatttcag	7200
aaagaatgct	gaccacaga	tggttagaga	ggcctacgca	gcaggctca	tcaagacgat	7260
ctaccggagt	aacaatctcc	aggagatcaa	atacctccc	aagaaggtta	aagatgcagt	7320
caaaagattc	aggactaatt	gcatcaagaa	cacagagaaa	gacatatttc	tcaagatcag	7380
aagtactatt	ccagatgga	cgattcaagg	cttgcttcat	aaaccaaggc	aagtaataga	7440
gattggagtc	tctaaaaagg	tagttcctac	tgaatctaag	gccatgcatg	gagtctaaga	7500
ttcaaatcga	ggatctaaca	gaactcgccg	tgaagactgg	cgaacagttc	atacagagtc	7560
ttttcagact	caatgacaag	aagaaaatct	tcgtcaacat	ggtaggagc	gacactctgg	7620
tctactccaa	aaatgtcaaa	gatacagtct	cagaagacca	aagggtatt	gagacttttc	7680
aacaaaggat	aatttcggga	aacctcctcg	gattccattg	cccagctatc	tgctacttca	7740
tcgaaaggac	agttagaaa	gaaggtggct	cctacaaatg	ccatcattgc	gataaaggaa	7800
aggtatcat	tcaagatgcc	tctgccgaca	gtggtccaa	agatggaccc	ccaccacga	7860
ggagcatcgt	gaaaaagaa	gacgttccaa	ccacgtcttc	aaagcaagtg	gattgatgtg	7920
acatctccac	tgacgtaagg	gatgacgcac	aatcccacta	tccttcgcaa	gacccttctc	7980
ctatataagg	aagttcattt	catttgagga	ggacacgctc	gagctcattt	ctctattact	8040
tcagccataa	caaaagaact	ctttctctct	cttattaaac	catgaaaaag	cctgaactca	8100
ccgcgacgtc	tgtcgagaag	tttctgatcg	aaaagttcga	cagcgtctcc	gacctgatgc	8160
agctctcggg	ggcgaaagaa	tctcgtgctt	tcagcttcga	tgtaggaggg	cgtggatatg	8220
tctcgggg	aaatagctgc	gccgatgggt	tctacaaaga	tcgttatggt	tatcggcact	8280
ttgcatcggc	cgctcctccg	attccggaag	tgcttgacat	tggggaattc	agcgagagcc	8340

-continued

---

tgacctattg	catctcccgc	cgtgcacagg	gtgtcacggt	gcaagacctg	cctgaaaccg	8400
aactgcccgc	tgttctgcag	ccggtcgcgg	aggccatgga	tgcgatcgct	gcggccgatc	8460
ttagccagac	gagcgggttc	ggcccattcg	gaccgcaagg	aatcggtcaa	tacactacat	8520
ggcgtgattt	catatgcgcg	attgctgata	cccatgtgta	tacttgcaa	actgtgatgg	8580
acgacaccgt	cagtgcgtcc	gtcgcgcagg	ctctcgatga	gctgatgctt	tgggccgagg	8640
actgcccga	agtcggcac	ctcgtgcacg	cggatttcgg	ctccaacaat	gtctgacgg	8700
acaatggcgg	cataacagcg	gtcattgact	ggagcgaggc	gatgttcggg	gattcccaat	8760
acgaggtcgc	caacatcttc	tcttgagggc	cgtggttggc	ttgtatggag	cagcagacgc	8820
gctacttcga	gcgaggcat	ccggagcttg	caggatcgcc	gcgctccgg	gcgtatatgc	8880
tccgcattgg	tcttgaccaa	ctctatcaga	gcttggttga	cggcaatttc	gatgatgcag	8940
cttggggcca	gggtcgatgc	gacgcaatcg	tccgatccgg	agccgggact	gtcgggcgta	9000
cacaaatcgc	ccgcagaagc	gcgcccgctc	ggaccgatgg	ctgtgtagaa	gtactcgccg	9060
atagtggaaa	ccgacgcccc	agcactcgtc	cgagggcaaa	ggaatagtga	ggtacctaaa	9120
gaaggagtgc	gtcgaagcag	atcgttcaaa	catttggcaa	taaagtttct	taagattgaa	9180
tctgttgcc	ggcttgcga	tgattatcat	ataatttctg	ttgaattacg	ttaagcatgt	9240
aataattaac	atgtaatgca	tgacgttatt	tatgagatgg	gtttttatga	ttagagtccc	9300
gcaattatac	atttaatacg	cgatagaaaa	caaaatatag	cgcgcaaac	aggataaatt	9360
atcgcgcgcg	gtgtcatcta	tgtaactaga	tcgatgtcga	atctgatcaa	cctgcattaa	9420
tgaatcggcc	aacgcgcggg	gagaggcggg	ttgcgtattg	ggcgtcttc	cgcttcctcg	9480
ctcactgact	cgctgcgctc	ggctcgttcg	ctgcggcgag	cggtatcagc	tactcaaaag	9540
gcggtaatc	ggttatccac	agaatcaggg	gataaacgcag	gaaagaacat	gtgagcaaaa	9600
ggccagcaaa	aggccaggaa	ccgtaaaaag	gccgcggttc	tggcgttttt	ccataggctc	9660
cgccccctg	acgagcatca	caaaaatcga	cgctcaagtc	agaggtggcg	aaacccgaca	9720
ggactataaa	gataaccaggc	gtttccccct	ggaagctccc	tcgtgcgctc	tctgttccg	9780
accctgccgc	ttaccggata	cctgtccgcc	tttctccctt	cgggaagcgt	ggcgctttct	9840
caatgctcac	gctgtaggta	tctcagttcg	gtgtaggtcg	tctcctcaa	gctgggctgt	9900
gtgcacgaac	cccccggtca	gcccgaaccg	tgcgccttat	ccggtaaacta	tcgtcttgag	9960
tccaaccggg	taagacacga	cttatcgcca	ctggcagcag	ccactggtaa	caggattagc	10020
agagcgaggt	atgtaggcgg	tgctacagag	ttcttgaagt	gggtggcctaa	ctacggctac	10080
actagaagga	cagtatttgg	tatctgcgct	ctgctgaagc	cagttacctt	cggaaaaaga	10140
gttggtagct	cttgatccgg	caaacaaacc	accgctggta	gcggtggttt	ttttgtttgc	10200
aagcagcaga	ttacgcgcag	aaaaaaagga	tctcaagaag	atcctttgat	cttttctacg	10260
gggtctgaog	ctcagtgga	cgaaaactca	cgtaaggga	ttttggtcat	gacattaacc	10320
tataaaaata	ggcgtatcac	gagcccttt	cgtctcgcgc	gtttcgggtga	tgacggtgaa	10380
aacctctgac	acatgcagct	cccggagacg	gtcacagctt	gtctgtaagc	ggatgcgggg	10440
agcagacaag	cccgtcaggg	cgcgtcagcg	ggtgttggcg	ggtgtcgggg	ctggcttaac	10500
tatgcggcat	cagagcagat	tgtactgaga	gtgcaccata	tggacatatt	gtcggttagaa	10560
cgcggtaca	attaatacat	aaccttatgt	atcatacaca	tacgatttag	gtgacactat	10620
agaacggcgc	gccaagcttt	tgatccatgc	ccttcatttg	ccgcttatta	attaatttgg	10680
taacagtcog	tactaatcag	ttacttatcc	ttccccatc	ataattaatc	ttgtagtct	10740

-continued

---

```

cgaatgccac aacctgact agtctcttgg atcataagaa aaagccaagg aacaaaagaa 10800
gacaaaacac aatgagagta tcctttgcat agcaatgtct aagttcataa aattcaaaca 10860
aaaacgcaat cacacacagt ggacatcact tatccactag ctgatcagga tcgccgcgctc 10920
aagaaaaaaaa aactggaccc caaaagccat gcacaacaac acgtactcac aaagggtgca 10980
atcgagcagc ccaaaacatt caccaactca acccatcatg agccctcaca tttgttgttt 11040
ctaaccacac ctcaaactcg tattctcttc cgccacctca tttttgttta tttcaacacc 11100
cgtcaaaactg catgccaccc cgtggccaaa tgtccatgca tgtaacaag acctatgact 11160
ataaatagct gcaatctcgg cccaggtttt catcatcaag aaccagttca atacctagct 11220
acaccgtatt aaagaattta agatatactg c 11251

```

```

<210> SEQ ID NO 7
<211> LENGTH: 9060
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: T-DNA of vector pZBL120xKS336
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: (1856)..(1856)
<223> OTHER INFORMATION: n is a, c, g, or t

```

```

<400> SEQUENCE: 7

```

```

aattacaacg gtatatatcc tgccgtcgac tctagaggat ccgcgcgctc gacggatata 60
atgagccgta aacaaagatg attaatagat aattaatacg tactagtaaa agtggcaaaa 120
gataacgaga aagaaccaat ttctttgcat tcggccttag cggaaggcat atataagctt 180
tgattatfff atttagtgta atgatttcgt acaaccaaag catttattta gtactctcac 240
acttgtgtcg cggccggccg ctacaggaac aggtgggggc ggccctcggc gcgctcgta 300
tgctccacga tgggtgtagc ctcggtgtgg gaggtgatgt ccagcttggg gtccacgtag 360
tagtagccgg gcagctgcac gggccttctg gccatgtaga tggacttgaa ctccaccagg 420
tagtggccgc cgtcctcag ctccagggcc ttgtggatct cgccctcag cacgcccgtc 480
cgggggtaca ggcgctcggg ggaggcctcc cagcccatag tcttctctcg cattacgggg 540
ccgtcggagg ggaagttcac gccgatgaac ttcacctgt agatgaagga gccgtcctgc 600
agggaggagt cctgggtcac ggtcaccaac ccgcccctc cgaagttcat cacgcccctc 660
cacttgaagc cctcggggaa ggacagcttc ttgtagtcgg ggatgtcggc ggggtgcttc 720
acgtacacct tggagccgta ctggaactgg ggggacagga tgtcccaggc gaagggcagg 780
gggcccctc tggcacctc cagcttggcg gtctgggtgc cctcgtaggg gcggcccctc 840
ccctcgcctc cgtatctgaa ctcgtggccg ttcacggagc cctccatgcg caccttgaag 900
cgcatgaact ccttgatgac gtccctcggg gaggccatgg gccgcttggg gggctatgga 960
agactttctt agttagtgt gtgaataagc aatgttggga gaatcgggac tacttatag 1020
ataggaataa aacagaaaag tattaagtgc taatgaaata tttagactga taattaaat 1080
cttcacgtat gtccacttga tataaaaaac tcaggaataa aggaagtaca gtagaattta 1140
aaggctactc ttttatatat acccgtgttc tctttttggc tagctagtgg cataaaaaat 1200
aatctatatt tttatcatta ttttaaatat cttatgagat ggtaaatatt tatcataatt 1260
tttttacta ttatttatta tttgtgtgtg taatacatat agaagttaat tacaaaat 1320
attactttt tcattatfff gatatgatc accattaatt tagtgttatt atttataata 1380

```

-continued

---

gttcatttta atctttttgt atatatatg cgtgcagtac ttttttctca catataacta	1440
ctattacatt ttatttatat aatattttta ttaatgaatt ttcgtgataa tatgtaatat	1500
tgttcattat tatttcagat tttttaaaaa tatttgggtt attatttatg aaatatgtaa	1560
tttttttagt atttgatttt atgatgataa agtgttctaa attcaaaaga agggggaaaag	1620
cgtaaacatt aaaaaacgtc atcaaaacaaa aacaaaatct tgtaataaaa gataaaactg	1680
tttgttttga tcaactgttat ttcgtaatat aaaaacatta tttatattta tattgttgac	1740
aaccaaattt gcctatcaaa tctaaccaat ataatgcatg cgtggcaggt aatgtactac	1800
catgaactta agtcatgaca taataaacgg tgaatctgac caatgcatgt acctanctaa	1860
attgtatttg tgacacgaag caaatgattc aattcacaat ggagatggga aacaaataat	1920
gaagaacca gaactaagaa agcttttctg aaaaataaaa taaaggcaat gtcaaaagta	1980
tactgcatca tcagtccaga aagcacatga tattttttta tcagtatcaa tgcagctagt	2040
tttattttac aatatcgata tagctagttt aaatatattg cagctagatt tataaatatt	2100
tgtgttatta tttatcattt gtgtaatcct gtttttagta ttttagttta tatatgatga	2160
taatgtattc caaatttaaa agaagggaaa taaattttaa caagaaaaaa agtcatcaaa	2220
caaaaaacaa atgaagggtt ggaagatgt taccatgtaa tgtgaatgtt acagtatttc	2280
ttttattata gagttaacaa attaactaat atgattttgt taataatgat aaaatatttt	2340
ttttattatt atttcataat ataaaaatag ttacttaat ataaaaaaaa ttctatcgtt	2400
cacaacaaag ttggccacct aatttaacca tgcatgtacc catggaccat attaggtaac	2460
catcaaacct gatgaagaga taaagagatg aagacttaag tcataacaca aaaccataaa	2520
aaacaaaaat acaatcaacc gtcaatctga ccaatgcatg aaaaagctgc aatagtgagt	2580
ggcgacacaa agcacatgat tttcttacia cggagataaa accaaaaaaa tatttcatga	2640
acaacctaga acaataaag cttttatata ataaatatat aaataaataa aggctatgga	2700
ataatatact tcaatatatt tggattaaat aaattgttgg cggggttgat atatttatac	2760
acacctaaag tcaactcaat ctcatcttca cttaactttt attttttttt tctttttatt	2820
tatcataaag agaattattga taataactt ttaacatat ttttatgaca tttttattg	2880
gtgaaaactt attaaaaatc ataaattttg taagttagat ttatttaaaag agttcctctt	2940
cttattttta attttttaat aaatttttaa ataactaaaa tttgtgttaa aaatgttaa	3000
aaagtgtgtt attaacctt ctcttcgagg atccaagctt gttgaaacat cctgaagtg	3060
tctcatttta ttttatttat tctttgctga taaaaaata aaataaaaga agctaagcac	3120
acggtcaacc attgctctac tgetaaaagg gttatgtgta gtgttttact gcataaatta	3180
tgcagcaaac aagacaactc aaattaaaa atttccttg cttgtttttt tgttgtctct	3240
gacttgactt tcttgggaa gttgggtgta taaggattgg gacaccattg tccttcttaa	3300
tttaatttta ttctttgctg ataaaaaaa aaatttcata tagtgtttaa taataatttg	3360
ttaaataacc aaaaagtcac atatgtttac tctcgtttaa ataattgaga ttcgtccagc	3420
aaggctaaac gattgtatag atttatgaca atatttactt ttttatagat aaatgttata	3480
ttataataaa tttatataca tatattatat gttatttatt attattttaa atccttcaat	3540
attttatcaa accaactcat aatttttttt ttatctgtaa gaagcaataa aattaaatag	3600
accacttta aggatgatcc aacctttata cagagtaaga gagttcaaat agtaccttt	3660
catatacata tcaactaaaa tattagaaat atcatggatc aaaccttata aagacattaa	3720
ataagtggat aagtataata tataaatggg tagtatataa tataaaatg gatcaaaact	3780

-continued

---

tctctcttta	taattgttat	gtctecttaa	catcctaata	taatacataa	gtgggtaata	3840
tataatata	aatggagac	aaactcttc	cattataatt	gttatgtctt	cttaacactt	3900
atgtctcgtt	cacaatgcta	aggttagaat	tgtttagaaa	gtcttatagt	acacatttgt	3960
ttttgtacta	tttgaagcat	tcataagcc	gtcacgattc	agatgattta	taataataag	4020
aggaaattta	tcatagaaca	ataaggtgca	tagatagagt	gttaatatat	cataacatcc	4080
tttgtttatt	catagaagaa	gtgagatgga	gctcagttat	tatactgtta	catggtcgga	4140
tacaatattc	catgctctcc	atgagctctt	acacctacat	gcatttttagt	tcataacttgc	4200
ggccgcgttg	cacacactga	tcacgctgg	aggttgtatc	ccctcttcca	tatccgaaag	4260
aacatccttc	agctggacat	cacctgcctc	aaacagactg	aaatctgcca	gacctacacc	4320
atccttgctt	ccttcgctag	ccgcagaaca	cccctcgaac	aggcatccga	tgttgcagtc	4380
gaagtctgog	tcagcgaacc	actcgctcag	attgattgag	ctgccc aaagt	ttggtctcga	4440
catggtgtct	agctcgtaat	ccatgcaagg	gctccacaag	ccctcttcca	tgctgtcgtc	4500
gactgtggtg	aggggctcog	cggtgtcgtc	aggcaccgog	ttctcatcgg	gttctgcaet	4560
gccatccggt	gtcttggett	cgcttgactc	cggtcttgc	tctgtagttc	cggtttcgtt	4620
ctgatcaggt	tgaggttcct	ggttgagtgc	cggcaccacc	tgtygctcct	gttggagctg	4680
tgccaggaac	agcgggtggt	ccaggtagca	gctgatgtcg	aagttggtta	cagcattgac	4740
gccacggtat	tcaatggccg	caaggtcata	ggccttggtt	gcctcttctt	gagtgtcaaa	4800
tgttcccaag	tagaggtaact	tgttcccaaa	gactcgccca	atccgtgctt	cccacctccc	4860
gttgtggtga	tgcttgccga	cgctctgta	cttgagagcg	ccctggaga	agccgctgct	4920
cctgcggcgg	agggaggcca	ggtactcttc	ccgggacacg	gcctccatct	ccggcatctc	4980
gctggagtaa	tcctccacag	ggaagttgag	cagagtctca	ggaccccagt	acttgagagc	5040
tgcgaggtea	taggcacgag	cagctgcctc	ctcgtgttca	tacgccccca	ggttagacttg	5100
cctgcctttc	ttcttgttgt	ggagcgggc	gaggcagtg	ttgtcccaga	gatgcgcttc	5160
gaaactgccc	gtccacctgt	gcttggtgac	tcccctgtag	acggagctcc	tcttcccgcg	5220
ggcggcggcg	gcggggtcct	tgccgatact	cttattaggc	tcgggcggcg	ccttgccgct	5280
cgccgccctc	cgctctttc	cgggaggac	gaggacggtg	tccgoggaga	cgaggacga	5340
ggaggaggac	ggcgacggtg	gcgagggaga	ctgcccgtga	gatctctcca	tgcgcccgca	5400
gtatatctta	aattctttaa	tacggtgtac	taggatattg	aactggttct	tgatgatgaa	5460
aacctgggcc	gagattgcag	ctatttatag	tcataggtct	tgtaacatg	catggacatt	5520
tggccacggg	gtggcatgca	gtttgacggg	tgttgaaata	aacaaaaatg	aggtggcgga	5580
agagaatacg	agtttgaggt	tgggttagaa	acaacaaatg	tgagggctca	tgatgggttg	5640
agttggtgaa	tgttttgggc	tgctcgattg	acacctttgt	gagtacgtgt	tgttgtgcat	5700
ggcttttggg	gtccagtttt	ttttcttga	cgccgcatc	ctgatcagct	agtggataag	5760
tgatgtccac	tgtgtgtgat	tgccgttttg	tttgaatttt	atgaacttag	acattgctat	5820
gcaaaggata	ctctcattgt	gttttgtctt	cttttgttcc	ttggcttttt	cttatgatcc	5880
aagagactag	tcagtgttgt	ggcattcgag	actaccaaga	ttaattatga	tgggggaagg	5940
ataagtaact	gattagtacg	gactgttacc	aaattaatta	ataagcggca	aatgaagggc	6000
atggatcaaa	agcttgggcg	gaattcactg	gccgtcgttt	tacaacgtcg	tgactgggaa	6060
aacctggcgg	ttaccaact	taatcgctt	gcagcacatc	ccccttctgc	cagctggcgt	6120

-continued

---

aatagcgaag aggcccgcac cgatcgccct tcccaacagt tgcgcagcct gaatggcgaa	6180
tggatcgatc cgtegatoga ccaagcggc catcgtgcct cccactcct gcagttcggg	6240
ggcatggatg cgcggatagc cgctgctggt ttctgggatg ccgacggatt tgcactgccg	6300
gtagaactcc gcgaggctgt ccagcctcag gcagcagctg aaccaactcg cgaggggatc	6360
gagccccgag tgagcctoga catggtgctg caaaattcgc cctggaccgg cccaacgatt	6420
tgtcgtcact gtcaagggtt gacctgcaact tcatttgggg ccacatata ccaaaaaaat	6480
gctgcataat tctcggggca gcaagtcggt taccggccg ccgtgctgga ccgggttga	6540
tgggtcccgat aactttcggg agagcggacg gccataactc aacttcaagg aatctcacc	6600
atgcccggcg gcggggaacc ggagttccct tcagtgaacg ttattagtgc gccgctcgg	6660
gtgtcgtaga tactagcccc tggggccttt tgaaattga ataagattta tgtaatcagt	6720
cttttaggtt tgaccgggtc tgcgctttt tttaaaattg gatttgaat aataaaacgc	6780
aattgtttgt tattgtggcg ctctatcata gatgtcgcta taaacctatt cagcacaata	6840
tattgttttc attttaatat tgtacatata agtagtaggg tacaatcagt aaattgaacg	6900
gagaatatta ttcataaaaa tacgatagta acgggtgata tattcattag aatgaaccga	6960
aaccggcggg aaggatctga gctacacatg ctccaggtttt ttacaacgtg cacaacagaa	7020
ttgaaagcaa atatcatgcg atcataggcg tctcgcatat ctcattaaag caggggggtg	7080
gcgaagaact ccagcatgag atccccgcgc tggaggatca tccagccggc gtcccggaaa	7140
acgattccga agcccaacct ttcatagaag gcggcgggtg aatcgaaatc tcgtgatggc	7200
aggttggggc tcgcttggtc ggtcatttcg aaccccagag tcccgtcag aagaactcgt	7260
caagaaggcg atagaaggcg atgcgctgag aatcgggagc ggcgataccg taaagcacga	7320
ggaagcggtc agcccattcg ccgccaagct ctccagcaat atcacgggta gccaacgcta	7380
tgtcctgata gcggtccgcc acaccagcc ggccacagtc gatgaatcca gaaaagcggc	7440
cattttccac catgatattc ggcaagcagg catcgccatg ggtaacgacg agatcctcgc	7500
cgtcgggcat gcgcgccttg agcctggcga acagttcggc tggcgcgagc ccctgatgct	7560
cttcgtccag atcatcctga tcgacaagac cggcttccat ccgagtaact gctcgtcga	7620
tgcgatgttt cgcttggtgg tcgaatgggc aggtagccgg atcaagcgtg tgcagccgcc	7680
gcattgcacg agccatgatg gatactttct cggcaggagc aaggtgagat gacaggagat	7740
cctgccccgg cacttcgccc aatagcagcc agtcccttcc cgcttcagtg acaacgtcga	7800
gcacagctgc gcaaggaacg cccgtcgtgg ccagccacga tagccgcgct gcctcgtcct	7860
gcagttcatt cagggcaacc gacaggtcgg tcttgacaaa aagaaccggg cgcctcgcg	7920
ctgacagcgg gaacacggcg gcatcagagc agccgattgt ctgttggtgccc cagtcatagc	7980
cgaatagcct ctccacccaa gcggccggag aacctgctg caatccatct tgttcaatca	8040
tgcgaaacga tccccgcaag cttggagact ggtgatttca gcgtgtcctc tccaaatgaa	8100
atgaacttcc ttatatagag gaagggtctt gcgaaggata gtgggattgt gcgtcatccc	8160
ttacgtcagt ggagatatca catcaatcca cttgcttga agacgtgggt ggaacgtctt	8220
ctttttccac gatgctcctc gtgggtgggg gtccatcttt gggaccactg tcggcagagg	8280
catcttcaac gatggccttt cctttatcgc aatgatggca tttgtaggag ccaccttctt	8340
tttccactat cttcacaata aagtgcagca tagctgggca atggaatccg aggaggttct	8400
cggatattac cctttgttga aaagtctcaa ttgccctttg gtcttctgag actgtatctt	8460
tgatattttt ggagtagaca agcgtgctgt gctccaccat gttgacgaag attttcttct	8520

-continued

---

```

tgtcattgag tcgtaagaga ctctgtatga actgttcgcc agtctttacg gcgagttctg 8580
ttaggtctctc tatttgaatc tttgactcca tggcctttga ttcagtggga actacctttt 8640
tagagactcc aatctctatt acttgcccttg gtttgtgaag caagccttga atcgccata 8700
ctggaatagt acttctgac ttgagaaata tatctttctc tgtgttcttg atgcagttag 8760
tcctgaatct tttgactgca tctttaacct tcttgggaag gtatttgatc tcctggagat 8820
tattgctcgg gtagatcgtc ttgatgagac ctgctgcgta agcctctcta accatctgtg 8880
ggttagcatt ctttctgaaa ttgaaaaggc taatcttctc attatcagtg gtgaacatgg 8940
tatcgtaacc ttctccgctg aacttcctga ctgatcgta gagatagagg aagtcgtcca 9000
ttgtgatctc tggggcaaaag gagatctgaa ttatcattta caattgaata taccctgcca 9060

```

```

<210> SEQ ID NO 8
<211> LENGTH: 34
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: PCR primer MWG339

```

```

<400> SEQUENCE: 8

```

```

gaattcgcgg cgcctatgag aaggctctcc tctg 34

```

```

<210> SEQ ID NO 9
<211> LENGTH: 34
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: PCR primer MWG340

```

```

<400> SEQUENCE: 9

```

```

gaattcgcgg cgcctcaaac cctaaattca cacg 34

```

```

<210> SEQ ID NO 10
<211> LENGTH: 11299
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: vector pKS333
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: (3588)..(3588)
<223> OTHER INFORMATION: n is a, c, g, or t

```

```

<400> SEQUENCE: 10

```

```

ggccgccatg agaaggtctc cctctgttct tacttctctc tcctctctct cctcctgcgt 60
cgccggcgcc ggcttcgaca gcaataatct caatctcgcc gccctctcgc gccggccgca 120
atcggagaag accggagcga aacgccggaa gcggaatcag gacgacgcca aatgcgagat 180
tgagaatcgt aacggtaata acaacaacag cagcaacaac aatgcctctt ccggccgccc 240
gagctccatt tacagaggag tcaactaggca ccgatggacc ggccggttcg aagcgcattc 300
ctgggacaag agttcgtgga atagcattca gaacaaaaaa ggaaggcaag tttatttggg 360
agcatacgat aacgaggaag ctgccgcccg aacttatgac ctgctgccc tcaagtactg 420
gggtcccgga accacctca atttcccggg agagtcgtac aggaatgaaa tagaagaaat 480
gcgaaaagt acgaaggagg agtatttggc gtcgttacgg cggcggagca gcggattttc 540
gagaggcgta tcgaagtacc gcggcgtggc ccgccaccac cacaacggcc ggtgggaggc 600
gcggatcgcc cgtgttttct gaagcaata tctttacctg ggaacttaca acacacaaga 660

```

-continued

---

ggaagcagca gcagcatatg acatggctgc aattgagtac agaggggtca atgcagtgac	720
caatttcgac atcagcaatt acattggcg gctggagaat aaatcatcag tttttccagc	780
agcagagcag ccctacagc ccaactgctc cctgcttcc tcttctgagg aaggcgaagt	840
agtacagcag caacagcaac agacgacgat ggcgttctca ggctgcccc tccagttccc	900
gtc gatggag aacagcccga cgacaatgga ggaggatcat gatctgcatt ggctattcct	960
agacacgggg ttcgtgcagg tccccgacct cccctcgag aagtctggcg aattgcctga	1020
cctgttcttt gatgagatcg ggttcgagga cgacatcggg ttgatattcg aggcgagctt	1080
ggaagacgag aggtgcgggg aggggggtga gaagttagaa gatgtggga aaatggagat	1140
gatgaagagt gatcatgagg agagggggtt gttctcgact acttgcctat cttcgtcgtc	1200
gataaccacc tcggtttcgt gtgaatttag ggtttgagcg gccgcaagta tgaactaaaa	1260
tgc atgtagg tgtaagagct catggagagc atggaatatt gtatccgacc atgtaacagt	1320
ataataactg agctccatct cacttcttct atgaataaac aaaggatgtt atgatatatt	1380
aacactctat ctatgcaoct tattgttcta tgataaattt cctcttatta ttataaatca	1440
tctgaatcgt gacggcttat ggaatgcttc aaatagtaca aaaacaaatg tgtactataa	1500
gactttctaa acaattctaa ccttagcatt gtgaacgaga cataagtgtt aagaagacat	1560
aacaattata atggaagaag tttgtctcca tttatatatt atatattacc cacttatgta	1620
ttatattagg atgttaagga gacataacaa ttataaagag agaagtttgt atccatttat	1680
atattatata ctaccattt atatattata cttatccact tatttaatgt ctttataagg	1740
tttgatccat gatatttcta atattttagt tgatagtat atgaaagggt actatttgaa	1800
ctctcttact ctgtataaag gttggatcat ccttaaagtg ggtctattta attttattgc	1860
ttcttacaga taaaaaaaa attatgagtt ggtttgataa aatattgaag gatttaaaat	1920
aataataaat aacatataat atagtatat aaatttatta taatataaca tttatctata	1980
aaaaagtaaa tattgtcata aatctataca atcgtttagc cttgctggac gaatctcaat	2040
tattttaaag agagtaaaaca tatttgactt tttggttatt taacaaatta ttatttaaca	2100
ctatatgaaa tttttttttt taccagcaaa gaataaaatt aaattaagaa ggacaatggt	2160
gtcccaatcc ttatacaacc aacttccaca agaaagtcaa gtcagagaca acaaaaaaac	2220
aagcaaaagga aattttttta tttgagttgt cttgtttgct gcataattta tgcagtaaaa	2280
cactacacat aaccctttta gcagtagagc aatggttgac cgtgtgotta gcttctttta	2340
ttttattttt ttatcagcaa agaataaata aaataaaatg agacacttca gggatgtttc	2400
aacaagcttg gatcctcgaa gagaaggggt aataacacac ttttttaaca tttttaacac	2460
aaattttagt tatttaaaaa tttattaaaa aatttaaaat aagaagagga actctttaa	2520
taaatctaac ttacaaaatt tatgattttt aataagtttt caccaataaa aaatgtcata	2580
aaaatagtgt aaaaagtata ttatcaatat tctctttatg ataataaaaa agaaaaaaaa	2640
aataaaagtt aagtgaaaat gagattgaag tgactttagg tgtgtataaa tatatcaacc	2700
ccgccaacaa tttatttaac ccaaatatat tgaagtatat tattccatag cttttattta	2760
tttatatatt tattatataa aagcctttatt tgttctaggt tggatcatgaa atattttttt	2820
ggttttatct ccgttgtaag aaaatcatgt gctttgtgct gccactcact attgcagctt	2880
tttcatgcat tggcagatt gacggttgat tgtatttttg ttttttatgg ttttgtgta	2940
tgacttaagt cttcatctct ttatctcttc atcaggtttg atggttacct aatatggtec	3000
atgggtacat gc atggttaa attaggtggc caactttggt gtgaacgata gaattttttt	3060

-continued

---

tatattaagt	aaactat	tttt	tatattatga	aataataata	aaaaa	atat	tttatcatta	3120
ttaacaaaat	catattagtt	aatttgtaa	ctctataata	aaagaaatac	tgtaacattc			3180
acattacatg	gtaacatcct	tccacccttt	catttggttt	ttgtttgatg	actttttttc			3240
ttgtttaaat	ttatttcoc	tcttttaaat	ttggaataca	ttatcatcat	atataaacta			3300
aaatactaaa	aacaggatta	cacaaatgat	aaataataac	acaaatattt	ataaatctag			3360
ctgcaatata	tttaactag	ctatatog	atgttaaat	aaaactagct	gcattgatac			3420
tgataaaaa	atatcatg	ctttctggac	tgatgatgca	gtatactttt	gacattgcct			3480
ttattttatt	tttcagaaaa	gctttcttag	ttctgggttc	ttcattattt	gtttcccatc			3540
tccattgtga	attgaatcat	ttgcttcgtg	tcacaaatac	aatttagnta	ggtacatgca			3600
ttggtcagat	tcacggttta	ttatgcatg	acttaagttc	atggtagtac	attacctgcc			3660
acgcatgcat	tatattgggt	agatttgata	ggcaaat	gttgcaaca	atataaatat			3720
aaataatggt	tttatattac	gaaataacag	tgatcaaac	aaacagtttt	atctttatta			3780
acaagat	ttttgtgtt	gatgacgtt	tttaatg	acgctttccc	ccttcttttg			3840
aatttagaac	actttatcat	cataaaatca	aatactaaaa	aaattacata	tttcataaat			3900
aataacacaa	atatttttaa	aaaatctgaa	ataataatga	acaatattac	atattatcac			3960
gaaaattcat	taataaaaat	attatataaa	taaaatgtaa	tagtagttat	atgtaggaaa			4020
aaagtactgc	acgcataata	tatacaaaaa	gattaanaatg	aactattata	aataataaca			4080
ctaaattaat	ggatgaatcat	atcaaaaaa	tgaaaaagta	aataaaat	gtaattaact			4140
tctatatgta	ttacacacac	aaataataaa	taatagtaaa	aaaaattatg	ataaatattt			4200
accatctcat	aagatattta	aaataatgat	aaaaatatag	attatttttt	atgcaactag			4260
ctagccaaaa	agagaacacg	ggtatatata	aaaagagtac	ctttaaatc	tactgtactt			4320
cctttattoc	tgacgttttt	atatcaagtg	gacatacgtg	aagattttta	ttatcagtct			4380
aaatatttca	ttagcactta	atacttttct	gttttattcc	tatcctataa	gtagtcccga			4440
ttctcccaac	attgcttatt	cacacaacta	actaagaaag	tcttccatag	ccccccaagc			4500
ggcccatggc	ctctcccgag	gacgtcatca	aggagtccat	gcgcttcaag	gtgcgcatgg			4560
agggctccgt	gaacggccac	gagttcgaga	tcgagggcga	ggcgaggggc	cgcccctacg			4620
agggcaccca	gaccgccaag	ctgaaggtga	ccaagggcgg	ccccctgcc	ttgcctggg			4680
acatcctgtc	ccccagttc	cagtacggct	ccaaggtgta	cgtgaagcac	cccgcgaca			4740
tccccgacta	caagaagctg	tccttccccg	agggttcaa	gtgggagcgc	gtgatgaact			4800
tcgaggacgg	cggcgtggg	accgtgacc	aggactcctc	cctgcaggac	ggetcctca			4860
tctacaaggt	gaagttcatc	ggcgtgaact	tcccctccga	cggccccgta	atgcagaaga			4920
agactatggg	ctgggaggcc	tccaccgagc	gcctgtaccc	ccgcgacggc	gtgctgaagg			4980
gcgagatcca	caaggccctg	aagctgaagg	acggcggcca	ctacctggtg	gagttcaagt			5040
ccatctacat	ggccaagaag	cccggtgcagc	tgcccggcta	ctactacgtg	gactccaagc			5100
tgacatcac	ctcccacaac	gaggactaca	ccatcgtgga	gcagtacgag	cgcgcgagg			5160
gccgccacca	cctgttctg	tagcggcccg	ccgcgacaca	agtgtagagag	tactaaataa			5220
atgctttgg	tgtacgaaat	cattacacta	aataaaataa	tcaaagctta	tatatgcctt			5280
ccgctaaggc	cgaatgcaaa	gaaattgggt	ctttctcgtt	atcttttgcc	acttttacta			5340
gtacgtatta	attactactt	aatcatcttt	gtttacggct	cattatatcc	gtcgaaggcg			5400

-continued

---

cgggcccgtc tagaactagt ggatccgtcg acggcgcgcc cgatcatccg gatatagttc	5460
ctcctttcag caaaaaaccc ctcaagaccc gtttagaggc cccaaggggt tatgctagtt	5520
attgctcagc ggtggcagca gccaaactcag cttectttcg ggctttgtta gcagccggat	5580
cgatccaagc tgtacctcac tattcctttg ccctcggacg agtgetgggg cgteggtttc	5640
cactatcggc gagtacttct acacagccat cggteccagac ggccgcgctt ctgcgggcga	5700
tttgtgtacg cccgacagtc ccggtcccg atcggacgat tgcgtcgcac cgaccctgcg	5760
cccaagctgc atcatcgaaa ttgccgtcaa ccaagctctg atagagttgg tcaagaccaa	5820
tgcggagcat atacgcccg agccgcggcg atcctgcaag ctccggatgc ctccgctcga	5880
agtagcgcgt ctgctgctcc atacaagcca accacggcct ccagaagaag atgttgcgga	5940
cctcgtattg ggaatccccg aacatcgct cgctccagtc aatgaccgct gttatgcggc	6000
cattgtccgt caggacattg ttggagccga aatccgcgtg cacgaggtgc cggacttcg	6060
ggcagtcctc ggcccaaaagc atcagctcat cgagagcctg cgcgacggac gcaactgacg	6120
tgtcgtccat cacagtttgc cagtgataca catggggatc agcaatcgcg catatgaaat	6180
cacgccatgt agtgatttga ccgattcctt gcggtccgaa tgggcccgaac ccgctcgtct	6240
ggctaagatc ggccgcagcg atcgcacca tagcctccgc gaccggctgc agaacagcgg	6300
gcagttcggg ttcaggcagg tcttgcaacg tgacaccctg tgcacggcgg gagatgcaat	6360
aggtcaggct ctcgctgaat tccccaatgt caagcacttc cggaaatcggg agcgcggccg	6420
atgcaaagtg ccgataaaca taacgatctt tgtagaaacc atcggcgcag ctatttacc	6480
gcaggacata tccacgcct cctacatcga agctgaaagc acgagattct tcgccctccg	6540
agagctgcat caggtcggag acgctgtcga acttttcgat cagaaacttc tcgacagacg	6600
tcgcggtgag ttcaggettt tccatgggta tatctccttc ttaaagttaa acaaaattat	6660
ttctagaggg aaaccgttgt ggtctcccta tagtgagtcg tattaatttc gcgggatcga	6720
gatcgatcca attccaatcc cacaaaaatc tgagcttaac agcacagttg ctctctcag	6780
agcagaatcg ggtattcaac accctcatat caactactac gttgtgtata acgggtccaca	6840
tgccggtata tacgatgact ggggttgtag aaaggcggca acaaacggcg ttcccggagt	6900
tgcacacaag aaatttgcca ctattacaga ggcaagagca gcagctgacg cgtacacaac	6960
aagttagcaa acagacaggt tgaacttcat ccccaaggga gaagctcaac tcaagcccaa	7020
gagctttgct aaggccctaa caagcccacc aaagcaaaaa gccactggc tcacgctagg	7080
aaccaaaagg cccagcagtg atccagcccc aaaagagatc tcctttgccc cggagattac	7140
aatggacgat ttctctatc tttacgatct aggaaggaag ttcgaagggtg aagggtgacga	7200
cactatgttc accactgata atgagaaggt tagcctcttc aatttcagaa agaatgctga	7260
cccacagatg gtttagagagg cctacgcagc aggtctcacc aagacgatct acccgagtaa	7320
caatctccag gagatcaaat accttcccaa gaaggttaaa gatgcagtca aaagattcag	7380
gactaattgc atcaagaaca cagagaaga catatttctc aagatcagaa gtactattcc	7440
agtatggacg attcaaggct tgcttcataa accaaggcaa gtaatagaga ttggagtctc	7500
taaaaaggta gttctactg aatctaaggc catgcatgga gtctaagatt caaatcgagg	7560
atctaacaga actcgcgctg aagactggcg aacagttcat acagagtctt ttacgactca	7620
atgacaagaa gaaaaacttc gtaacatgg tggagcacga cactctggtc tactccaaaa	7680
atgtcaaaga tacagctcga gaagaccaa gggctattga gacttttcaa caaaggataa	7740
tttcgggaaa cctcctcggg ttccattgcc cagctatctg tcaactcacc gaaaggacag	7800

-continued

---

tagaaaagga	aggtggctcc	tacaaatgcc	atcattgcca	taaaggaaag	gctatcattc	7860
aagatgcctc	tgccgacagt	ggccccaaag	atggaccccc	accacagagg	agcatcgtgg	7920
aaaaagaaga	cgttccaacc	acgtcttcaa	agcaagtggg	ttgatgtgac	atctccactg	7980
acgtaagggg	tgacgcacaa	tcccactatc	cttcgcaaga	cccttcctct	atataaggaa	8040
gttcatttca	tttgagagg	acacgctcga	gctcatttct	ctattacttc	agccataaca	8100
aaagaactct	tttctcttct	tattaaacca	tgaaaaagcc	tgaactcacc	gcgacgtctg	8160
tcgagaagtt	tctgatcgaa	aagtctgaca	gcgtctccga	cctgatgcag	ctctcggagg	8220
gcgaagaatc	tcgtgcttcc	agcttcgatg	taggagggcg	tggatatgtc	ctgcgggtaa	8280
atagctgcgc	cgatggtttc	tacaaagatc	gttatgttta	tcggcacttt	gcatcggccc	8340
cgctcccgat	tccggaagtg	cttgacattg	gggaattcag	cgagagcctg	acctattgca	8400
tctccgcgcg	tgacacaggt	gtcacgttgc	aagacctgcc	tgaaaccgaa	ctgcccgctg	8460
ttctgcagcc	ggctgcggag	gccatggatg	cgatcgctgc	ggccgatctt	agccagacga	8520
gcggttctcg	cccattcggg	ccgcaaggaa	tcggtcaata	cactacatgg	cgatgattca	8580
tatgcgcgat	tgctgatccc	catgtgtatc	actggcaaac	tgtgatggac	gacaccgtca	8640
gtgcgtccgt	cgcgcaggct	ctcgatgagc	tgatgctttg	ggccgaggac	tgccccgaag	8700
tccggcacct	cgtgcacgcg	gatttcggct	ccaacaatgt	cctgacggac	aatggccgca	8760
taacagcggg	cattgactgg	agcagggcga	tgctcgggga	ttccaatac	gaggtcgcca	8820
acatcttctt	ctggaggccg	tggttggtct	gtatggagca	gcagacgcgc	tacttcgagc	8880
ggaggcatcc	ggagcttgca	ggatcgccgc	ggctccgggc	gtatatgtct	cgcatgtgtc	8940
ttgaccaact	ctatcagagc	ttggttgacg	gcaatttcga	tgatgcagct	tgggcgcagg	9000
gtcgatgcga	cgcaatcgtc	cgatccggag	ccgggactgt	cgggcgtaca	caaatcgccc	9060
gcagaagcgc	ggcgcgtctg	accgatggct	gtgtagaagt	actcgcgat	agtggaaacc	9120
gacgccccag	cactcgtccg	agggcaaagg	aatagttagg	tacctaaaga	aggagtgcgt	9180
cgaagcagat	cgttcaaaaca	ttgggcaata	aagtttctta	agattgaatc	ctggttccgg	9240
tcttgcgatg	attatcatat	aatttctgtt	gaattacggt	aagcatgtaa	taattaacat	9300
gtaatgcatg	acgttattta	tgagatgggt	ttttatgatt	agagtcccgc	aattatacat	9360
ttaatacgcg	atagaaaaaca	aaatatagcg	cgcaaaactag	gataaattat	cgcgcgcggg	9420
gtcatctatg	ttactagatc	gatgtcgaat	ctgatcaacc	tgcattaatg	aatcgcccaa	9480
cgcgcggggg	gaggcgggtt	gcgtattggg	cgctcttccg	cttcctcget	cactgaactc	9540
ctgcgctcgg	tcgttcggct	gcggcgagcg	gtatcagctc	actcaaaggc	ggtaatacgg	9600
ttatccacag	aatcagggga	taacgcagga	aagaacatgt	gagcaaaagg	ccagcaaaag	9660
gccaggaacc	gtaaaaaggc	cgcgttgctg	gcgtttttcc	ataggctccg	ccccctgac	9720
gagcatcaca	aaaatcgacg	ctcaagtccg	aggtggcgaa	acccgacagg	actataaaga	9780
taccagcgct	tccccctgg	aagctccctc	gtgcgctctc	ctggtccgac	cctgcgcgct	9840
accggatacc	tgtccgcctt	tctcccttcg	ggaagcgtgg	cgctttctca	atgctcacgc	9900
tgtaggtatc	tcagttcggg	gtaggtcgtt	cgctccaagc	tgggctgtgt	gcacgaaccc	9960
cccgttcagc	ccgaccgctg	cgccttatcc	ggtaactatc	gtcttgagtc	caaccgggta	10020
agacacgact	tatcgccact	ggcagcagcc	actggtaaca	ggattagcag	agcgagggat	10080
gtaggcgggt	ctacagagtt	cttgaagtgg	tggcctaact	acggctacac	tagaaggaca	10140

-continued

---

```

gtatttgga tctgcgctct gctgaagcca gttaccttcg gaaaaagagt tggtagctct 10200
tgatccggca aaaaaccac cgctggtagc ggtgggtttt ttgtttgcaa gcagcagatt 10260
acgcgcagaa aaaaaggatc tcaagaagat cctttgatct tttctacggg gtctgacgct 10320
cagtggaaoc aaaactcacg ttaagggatt ttggatcatg cattaaccta taaaaatagg 10380
cgtatcacga ggccctttcg tctcgcgctt ttcgggtgatg acggtgaaaa cctctgacac 10440
atgcagctcc cggagacggt cacagcttgt ctgtaagcgg atgccgggag cagacaagcc 10500
cgtcagggcg cgtcagcggg tgttggcggg tgctcgggct ggcttaacta tgcggcatca 10560
gagcagattg tactgagagt gcaccatagc gacatattgt cgtagaacg cggctacaat 10620
taatacataa ccttatgtat catacacata cgatttaggt gacactatag aacggcgcgc 10680
caagcttttg atccatgccc ttcatttgcc gcttattaat taatttgga acagtccgta 10740
ctaactcagtt acttatecct cccccatcat aattaactct ggtagtctcg aatgccacaa 10800
cactgactag tctcttggat cataagaaaa agccaaggaa caaagaaga caaacacaaa 10860
tgagagtatc ctttgcatac caatgtctaa gttcataaaa ttcaacaaa aacgcaatca 10920
cacacagtgg acatcactta tccactagct gatcaggatc gccgcgctca gaaaaaaaa 10980
ctggacccca aaagccatgc acaacaacac gtactcaca aggtgtcaat cgagcagccc 11040
aaaaattca ccaactcaac ccatcatgag cctcactatt tgggtttct aaccaacct 11100
caactcgtga ttctcttcgc ccacctcatt tttgtttatt tcaacacccg tcaactgca 11160
tgccaccccg tggccaaatg tccatgcatg ttaacaagac ctatgactat aaatagctgc 11220
aatctcggcc caggttttca tcatcaagaa ccagttcaat atcctagtag accgtattaa 11280
agaatttaag atatactgc 11299

```

```

<210> SEQ ID NO 11
<211> LENGTH: 9142
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: T-DNA of vector pZBL120xKS333
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: (1892)..(1892)
<223> OTHER INFORMATION: n is a, c, g, or t

```

```

<400> SEQUENCE: 11

```

```

aattacaacg gtatatatcc tgccgtcgac tctagaggat ccgcgcccgc gacggatcca 60
ctagtcttag agcggcccgc gccgtcgacg gatataatga gccgtaaca aagatgatta 120
agtagtaatt aatacgtact agtaaaagtg gcaaaagata acgagaaaga accaatttct 180
ttgcattcgg ccttagcggg aggcataatg aagctttgat tattttattt agtgtaatga 240
ttctgtacaa ccaaaagcatt tatttagtac tctcacactt gtgtcgcggc cggccgctac 300
aggaacaggt ggtggcggcc ctcggcgcgc tcgtactgct ccacgatggt gtagtctctc 360
ttgtgggagg tgatgtccag cttggagtcc acgtagtagt agccgggagc ctgcacgggc 420
ttcttgccca tgtagatgga cttgaaactc accagtagt ggccgcccgc cttcagcttc 480
agggccttgt ggatctcgcc cttcagcacg ccgtcgcggg ggtacaggcg ctcggtggag 540
gcctcccagc ccatagtctt cttctgcatt acggggcctg cggaggggaa gttcacgccc 600
atgaacttca cctttagatg gaaggagccg tctcgcaggg aggagtcctg ggtcacggtc 660
accacgcccg cgtcctcgaa gttcatcacg cgtcctccact tgaagccctc ggggaaggac 720
agcttcttgt agtcggggat gtcggcgggg tgcttcacgt acaccttggg gccgtactgg 780

```

-continued

---

aactgggggg	acaggatgtc	ccaggcgaag	ggcagggggc	cgccttggg	caccttcagc	840
ttggcggctc	gggtgccctc	gtaggggccc	ccctcgcctc	cgccttcgat	ctcgaactcg	900
tggccgttca	cggagccctc	catgcgcacc	ttgaagcgca	tgaactcctt	gatgacgtcc	960
tggaggagg	ccatgggccc	cttggggggc	tatggaagac	tttcttagtt	agttgtgtga	1020
ataagcaatg	ttgggagaat	cgggactact	tataggatag	gaataaaaca	gaaaagtatt	1080
aagtgcta	aatatttta	gactgataat	taaaatcttc	acgtatgtcc	acttgatata	1140
aaaacgtcag	gaataaagga	agtacagtag	aatttaaagg	tactcttttt	atatataccc	1200
gtgttctctt	tttggctagc	tagttgcata	aaaaataatc	tatatcttta	tcattatctt	1260
aaatatctta	tgagatggta	aatattttatc	ataatttttt	ttactattat	ttattatttg	1320
tgtgtgta	acatatagaa	gttaattaca	aattttatct	actttttcat	tattttgata	1380
tgattcacca	ttaatttagt	gttattatct	ataatagttc	attttaatct	tttgtatat	1440
attatgcgtg	cagtactttt	ttcctacata	taactactat	tacattttat	ttatataata	1500
tttttattaa	tgaattttcg	tgataaatag	taataattgtt	cattattatt	tcagattttt	1560
taaaaatatt	tgtgttatta	tttatgaaat	atgtaatttt	tttagtattt	gattttatga	1620
tgataaagtg	ttctaatttc	aaaagaaggg	ggaaagcgta	aacattaaaa	aacgtcatca	1680
aacaaaaaca	aatcttgggt	aataaagata	aaactgtttg	ttttgatcac	tgttattttcg	1740
taataataaaa	acattattta	tatttatatt	gttgacaacc	aaatttgcc	atcaaatcta	1800
accaatataa	tgcatgcgtg	gcaggtaatg	tactaccatg	aacttaagtc	atgacataat	1860
aaaccgtgaa	tctgaccaat	gcactgtacct	anctaaattg	tattttgtgac	acgaagcaaa	1920
tgattcaatt	cacaatggag	atgggaaaca	aataatgaag	aaccagaac	taagaaagct	1980
tttctgaaaa	ataaaataaa	ggcaatgtca	aaagtatact	gcacatcag	tccagaaagc	2040
acatgatatt	ttttatcag	tatcaatgca	gctagtctta	ttttacaata	tcgatatagc	2100
tagtttaaat	atattgcagc	tagatttata	aatattttgtg	ttattattta	tcattttgtg	2160
aatctgtttt	ttagtatttt	agtttatata	tgatgataat	gtattccaaa	tttaaagaa	2220
gggaaataaa	tttaacaag	aaaaaaagtc	atcaaacaaa	aaacaaatga	aagggtggaa	2280
agatgttacc	atgtaatgtg	aatgttacag	tatttctttt	attatagagt	taacaaatta	2340
actaatatga	ttttgtta	aatgataaaa	tatttttttt	attattattt	cataatataa	2400
aaatagttta	cttaataataa	aaaaaattct	atcgttcaca	acaaagttgg	ccacctta	2460
taacctatgca	tgtacctatg	gacctatata	ggtaacctatc	aaacctgatg	aagagataaa	2520
gagatgaaga	cttaagtc	aacacaaaac	cataaaaaac	aaaaatacaa	tcaacctgca	2580
atctgaccaa	tgcatgaaaa	agctgcaata	gtgagtgccg	acacaaagca	catgattttc	2640
ttacaacgga	gataaaacca	aaaaaatatt	tcatgaacaa	cctagaacaa	ataaagcttt	2700
tataataataa	atatataaat	aaataaagc	tatggaataa	tatacttcaa	tatatttgga	2760
ttaaataaat	tgttggcggg	gttgatata	ttatacacac	ctaaagtcac	ttcaatctca	2820
ttttcactta	acttttattt	ttttttctt	tttattttatc	ataaagagaa	tattgataat	2880
atacttttta	acatattttt	atgacatttt	ttattgttga	aaacttatta	aaaaatcataa	2940
atatttgaag	ttagatttat	ttaaagagtt	cctcttctta	ttttaaattt	tttaataaat	3000
ttttaaataa	ctaaaatttg	tgtaaaaaat	gttaaaaaag	tgtgttatta	acccttctct	3060
tcgaggatcc	aagcttgggtg	aaacatccct	gaagtgtctc	attttatttt	atttattctt	3120

-continued

---

tgctgataaa	aaaataaaat	aaaagaagct	aagcacacgg	tcaaccattg	ctctactgct	3180
aaaaggggta	tgtgtagtgt	tttactgcat	aaattatgca	gcaaacaaga	caactcaaat	3240
taaaaaat	cctttgcttg	ttttttgtt	gtctctgact	tgactttcct	gtggaagtgt	3300
gttgtataag	gattgggaca	ccattgtcct	tcttaattta	atthttattct	ttgctgataa	3360
aaaaaaaaat	ttcatatagt	gttaaataat	aatttggtta	ataacaaaa	agtcaaatat	3420
gtttactctc	gtttaaataa	ttgagattcg	tccagcaagg	ctaaacgatt	gtatagattt	3480
atgacaatat	ttactttttt	atagataaat	gttatattat	aataaattta	tatacatata	3540
ttatatgtta	tttattatta	ttttaaatac	ttcaatattt	tatcaaacca	actcataatt	3600
ttttttttat	ctgtaagaag	caataaaatt	aaatagacc	actttaagga	tgatccaacc	3660
tttatacaga	gtaagagagt	tcaaatagta	ccctttcata	tacatatcaa	ctaaaatatt	3720
agaaatatca	tggatcaaac	cttataaaga	cattaataaa	gtggataagt	ataatatata	3780
aatgggtagt	atataatata	taaattggata	caaacttctc	tctttataat	tgttatgtct	3840
ccttaacatc	ctaataat	acataagtgg	gtaatatata	atatataaat	ggagacaaac	3900
ttcttcatt	ataattgtta	tgtctcttta	acacttatgt	ctcgttcaca	atgctaaggt	3960
tagaattggt	tagaaagtct	tatagtacac	atthgttttt	gtactatttg	aagcattcca	4020
taagccgtca	cgattcagat	gatttataat	aataagagga	aatttatcat	agaacaataa	4080
ggtgcataga	tagagtgtta	atatatcata	acatcctttg	tttattcata	gaagaagtga	4140
gatggagctc	agttattata	ctgttacatg	gtcggataca	atattccatg	ctctccatga	4200
gctcttacac	ctacatgat	tttagttcat	acttgoggcc	gctcaaacc	taaattcaca	4260
cgaaaccgag	gtggttatcg	acgacgaaga	tggcgaagta	gtcgagaaca	acccctctc	4320
ctcatgatca	ctcttcac	tctccatttt	ccccacatct	tctaacttct	cacccctctc	4380
cccgcacctc	tcgtcttcca	agctcgctc	gaatatcaac	ccgatgtcgt	cctcgaacc	4440
gatctcatca	aagaacaggt	caggcaatc	gccagacttc	tcgaggggga	ggtcggggac	4500
ctgcacgaac	cccgtgtcta	ggaatgacca	atgcagatca	tgatcctcct	ccattgtcgt	4560
cgggctgttc	tccatcgacg	ggaactggag	gggcgagcct	gagaacgcca	tcgtcgtctg	4620
ttgctgttgc	tgtctgacta	cttcgccttc	ctcagaagag	gaagcagggg	agcagttggg	4680
ctgtaggggc	tgtctcgtctg	ctggaaaaac	tgatgattta	ttctccagcc	gcccattgta	4740
attgctgatg	tcgaaattgg	tactgcatt	gacccctctg	tactcaattg	cagccatgtc	4800
atatgctgct	gctgcttct	cttgtgtgtt	gtaagttccc	aggtaaagat	atttgcctcc	4860
gaaaaacacg	ccgatccg	cctcccaccg	gccgttgtgg	tgggtggcggg	ccacgccgcg	4920
gtacttcgat	acgcctctcg	aaaatccgct	gctccgcgcg	cgtaacgacg	ccaaactctc	4980
ctccttcgta	actttccgca	ttctctctat	ttcattcctg	taegactcta	ccgggaaatt	5040
gaggggtggt	ccgggacccc	agtacttgag	ggcagcgagg	tcataagttc	ggggcgcagc	5100
ttcctcgta	tcgtatgctc	ccaaataaac	ttgccttct	ttttgttct	gaatgctatt	5160
ccacgaactc	ttgtcccaga	gatgcgcttc	gaaccggccg	gtccatcggg	gcctagtgac	5220
tcctctgtaa	atggagctcc	ggcggccgga	agaggcattg	ttgttgetgc	tgttgttgtt	5280
attaccgtta	cgattctcaa	tctcgcattt	ggcgtcgtcc	tgattccgct	tcggcggttt	5340
cgctccggtc	ttctccgatt	gcggccggcg	cggagggggc	gcgagattga	gattattgct	5400
gtcgaagccg	ccgccgcca	cgcaggagga	ggaggaggag	gaggaagtag	aaacagaggg	5460
agaccttctc	atggcggccg	cagtatatct	taaattcttt	aatacggtgt	actaggatat	5520

-continued

---

tgaactgggt	cttgatgatg	aaaacctggg	ccgagattgc	agctatttat	agtcataggt	5580
cttgtaaca	tgcatggaca	ttggccacg	gggtggcatg	cagtttgacg	ggtggtgaaa	5640
taacaaaaa	tgaggtggcg	gaagagaata	cgagtttgag	gttgggttag	aaacaacaaa	5700
tgtgagggt	catgatgggt	tgagttggg	aatgttttg	gctgctcgat	tgacacctt	5760
gtgagtacgt	gttgttgtc	atggctttg	gggtccagtt	ttttttctt	gacgcggcga	5820
tcctgatcag	ctagtggata	agtgatgtcc	actgtgtgtg	attgcgtttt	tgtttgaatt	5880
ttatgaact	agacattgct	atgcaaagga	tactctcatt	gtgtttgtc	ttctttgtt	5940
ccttggctt	ttcttatgat	ccaagagact	agtcagtggt	gtggcattcg	agactaccaa	6000
gattaattat	gatgggggaa	ggataagtaa	ctgattagta	cggactgtta	ccaaattaat	6060
taataagcgg	caaatgaagg	gcattggatca	aaagcttggc	gcgaattcac	tgcccgctgt	6120
tttacaacgt	cgtgactggg	aaaaccctgg	cgtaaccaa	cttaatcgcc	ttgcagcaca	6180
tcccccttc	gccagctggc	gtaatagcga	agaggcccg	accgatcgcc	cttccaaca	6240
gttgccgagc	ctgaatggcg	aatggatcga	tccgtcgatc	gaccaaagcg	gccatcggtc	6300
ctccccactc	ctgcagttcg	ggggcatgga	tgccgggata	gccgctgctg	gttctctgga	6360
tgccgacgga	tttgactgct	cggtagaact	ccgagaggtc	gtccagcctc	aggcagcagc	6420
tgaaccaact	cgccagggga	tcgagcccct	gctgagcctc	gacatgttgt	cgcaaaattc	6480
gccctggacc	cgcccaacga	ttgtcgtca	ctgtcaaggt	ttgacctgca	cttcatttgg	6540
ggcccacata	cacaaaaaaa	atgctgcata	attctcgggg	cagcaagtcg	gttaccggc	6600
cgccgtgctg	gaccgggttg	aatggtgcc	gtaactttcg	gtagagcgga	cgccaatac	6660
tcaacttcaa	ggaatctcac	ccatgcgcgc	cgccggggaa	ccggagttcc	cttcagttaa	6720
cgttattagt	tcgcccctcg	gtgtgtcgt	gatactagcc	cctggggcct	tttgaattt	6780
gaataagatt	tatgtaatca	gtcttttagg	tttgaccggt	tctgcccgtt	tttttaaat	6840
tggatttgta	ataataaac	gcaattgtt	gttattgtgg	cgctctatca	tagatgtcgc	6900
tataaaccta	ttcagcaca	tatattgtt	tcattttaat	attgtacata	taagtagtag	6960
ggtacaatca	gtaaatgaa	cgagagaata	tattcataaa	aatacgatag	taaccgggtga	7020
tatattcatt	agaatgaacc	gaaaccggcg	gtaaggatct	gagctacaca	tgctcaggtt	7080
ttttacaacg	tgcaaacag	aattgaaagc	aaatatcatg	cgatcatagg	cgctcgcgat	7140
atctcattaa	agcagggggg	gggcgaagaa	ctccagcatg	agatccccgc	gctggaggat	7200
catccagccg	gctccccgga	aaacgattcc	gaagcccaac	ctttcataga	aggcggcggg	7260
ggaatcga	tctcgtgatg	gcaggttggg	cgctcgttgg	tcggtcattt	cgaaccccg	7320
agtcccgtc	agaagaactc	gtcaagaagg	cgatagaagg	cgatgcgctg	cgaatcggga	7380
gcggcgat	cgtaaacgac	gaggaagcgg	tcagcccatt	cgccgccaag	ctcttcagca	7440
atatcacggg	tagccaacgc	tatgtcctga	tagcgggtccg	ccacaccag	ccggccacag	7500
tcgatgaatc	cagaaaagcg	gccattttcc	accatgat	tcggcaagca	ggcatcgcca	7560
tgggtcacga	cgagatcctc	gccgtcgggc	atgcgcgct	tgagcctggc	gaacagttcg	7620
gctggcgcga	gccccctgatg	ctcttcgtcc	agatcatcct	gatcgacaag	accggttcc	7680
atccgagtac	gtgctcgtc	gatgcgatgt	ttcgcttggg	ggtcgaatgg	gcaggtagcc	7740
ggatcaagcg	tatgcagccg	ccgcattgca	tcagccatga	tggatacttt	ctcggcagga	7800
gcaaggtgag	atgacaggag	atcctgcccc	ggcacttcgc	ccaatagcag	ccagtcctt	7860

-continued

---

```

cccgcttcag tgacaacgtc gagcacagct gcgcaaggaa cgcccgtcgt ggccagccac 7920
gatagccgcg ctgcctcgtc ctgcagttca ttcagggcac cggacaggtc ggtcttgaca 7980
aaaagaaccg ggcgccctg cgctgacagc cggaaacacgg cggcatcaga gcagccgatt 8040
gtctgtttg  cccagtcata gccgaatagc ctctccacc aagcggccgg agaacctgcg 8100
tgcaatccat cttgttcaat catgcgaaac gatccccgca agcttgagaga ctggtgattt 8160
cagcgtgtcc tctccaaatg aaatgaactt ccttatatag aggaagggtc ttgcgaagga 8220
tagtgggatt gtgcgtcadc ccttacgtca gtggagatat cacatcaatc cacttgcttt 8280
gaagacgtgg ttggaacgtc ttctttttcc acgatgctcc tcgtgggtgg ggggccatct 8340
ttgggaccac tgtcggcaga ggcattctca acgatggcct ttcctttatc gcaatgatgg 8400
catttgtagg agccaccttc cttttccact atcttcacaa taaagtgaca gatagctggg 8460
caatggaatc cgaggagggt tccggatatt accctttggt gaaaagtctc aattgcctt 8520
tggctctctg agactgtatc tttgatattt ttggagtaga caagcgtgtc gtgctccacc 8580
atggtgacga agattttctt cttgtcattg agtcgtaaga gactctgtat gaactgttcg 8640
ccagtcctta cggcgagttc tgttaggtcc tctattttaa tctttgactc catggccttt 8700
gattcagtg  gaactacctt tttagagact ccaatcteta ttacttgctt tggtttgta 8760
agcaagcctt gaatcgtoca tactggaata gtacttctga tcttgagaaa tatactttc 8820
tctgtgttct tgatgcagtt agtcctgaat cttttgactg catctttaac cttcttgga 8880
aggatattga tctcctggag attattgctc gggtagatcg tcttgatgag acctgctgcg 8940
taagcctctc taaccatctg tgggttagca ttctttctga aattgaaaag gctaactctc 9000
tcattatcag tgggtaacat ggtatcgtca ccttctcctg cgaacttctc gactagatcg 9060
tagagataga ggaagtcgtc cattgtgatc tctggggcaa aggagatctg aattatcatt 9120
tacaattgaa tatatcctgc ca 9142

```

```

<210> SEQ ID NO 12
<211> LENGTH: 34
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: PCR primer MWG341

```

```
<400> SEQUENCE: 12
```

```
gaattcgcgg ccgcatgaag aggtctccag catc 34
```

```

<210> SEQ ID NO 13
<211> LENGTH: 34
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: PCR primer MWG342

```

```
<400> SEQUENCE: 13
```

```
gaattcgcgg ccgctcatag atctagagca tagt 34
```

```

<210> SEQ ID NO 14
<211> LENGTH: 11304
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: vector pKS334
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: (3593)..(3593)
<223> OTHER INFORMATION: n is a, c, g, or t

```

-continued

&lt;400&gt; SEQUENCE: 14

ggccgcatga agaggtctcc agcatcttct tgttcacat ctacttctc tgttgggtt 60  
 gaagctccca ttgaaaaag aaggcctaag catccaagga ggaataattt gaagtcacaa 120  
 aatgcaagc agaaccaaac caccactggt gccagaagaa gctctatcta tagaggagt 180  
 acaaggcata ggtggacagg gaggttgaa gctcacctat gggataagag ctcttggaac 240  
 aacattcaga gcaagaagg tgcacaagt tatttggggg catatgatac tgaagaatct 300  
 gcagcccgta cctatgaact tgcagccctt aaatactggg gaaaagatgc aaccctgaat 360  
 ttcccgatag aaacttatac caaggagctc gaggaatgg acaaggtttc aagagaagaa 420  
 tatttggctt ctttgcggcg ccaaacgagt ggcttttcta gaggcctgtc taagtaccgt 480  
 ggggttgcta ggcacatca taatggtcgc tgggaagcac gaattggaag agtatgcgga 540  
 aacaagtaoc tctacttggg gacatataaa actcaagagg aggcagcagt ggcataatgac 600  
 atggcagcaa tagagtaccg tggagtcaat gcagtgacca attttgacat aagcaactac 660  
 atggacaaaa taaagaagaa aaatgaccaa acccaacaac aacaacaga agcacaacg 720  
 gaaacagttc ctaactctc tgactctgaa gaagtagaag tagaacaaca gacaacaaca 780  
 ataaccacac caccctcctc tgaaaatctg cacatgccac cacagcagca ccaagttaa 840  
 tacaccccc atgtctctcc aaggaagaa gaatcatcat cactgatcac aattatggac 900  
 catgtgcttg agcaggatct gccatggagc ttcattgata ctggcttgtc tcagttcaa 960  
 gatccaaact tggcttctc caaaggatg gatgacttgg tggcagatgt tgatagtgca 1020  
 ggggttgagg aagacattga tttctgttc agcactcaac ctggatgata gactgagagt 1080  
 gatgtcaaca atatgagcgc agttttgat agtgttgagt gtggagacac aaatggggct 1140  
 ggtggaagca tgatgcatgt ggataacaag cagaagatag tatcatttgc ttcttccca 1200  
 tcatctacaa ctacagtttc ttgtgactat gctctagatc tagcggcgc aagtatgac 1260  
 taaaatgcat gtaggtgtaa gagctcatgg agagcatgga atattgtatc cgaccatgta 1320  
 acagtataat aactgagctc catctcactt cttctatgaa taaacaaagg atgttatgat 1380  
 atattaacac tctatctatg caccttattg ttctatgata aatttctctt tattattata 1440  
 aatcatctga atcgtgacgg cttatggaat gcttcaata gtacaaaaac aaatgtgtac 1500  
 tataagactt tctaacaat tctaacctta gcattgtgaa cgagacataa gtgttaagaa 1560  
 gacataacaa ttataatgga agaagtttgt ctccatttat atattatata ttaccactt 1620  
 atgtattata ttaggatgtt aaggagacat aacaattata aagagagaag tttgatcca 1680  
 tttatataat atatactacc catttatata ttatacttat ccacttattt aatgtcttta 1740  
 taaggtttga tccatgatat ttctaataat ttagttgata tgtatatgaa aggggtactat 1800  
 ttgaactctc ttactctgta taaaggttgg atcatcctta aagtgggtct atttaatttt 1860  
 attgcttctt acagataaaa aaaaaattat gagttggttt gataaaatat tgaaggattt 1920  
 aaaataataa taaataacat ataatatatg tatataaatt tattataata taacatttat 1980  
 ctataaaaaa gtaaatattg tcataaatct atacaatcgt ttagccttgc tggacgaatc 2040  
 tcaattattt aaacgagagt aaacatattt gactttttgg ttatttaaca aattattatt 2100  
 taacactata tgaaattttt tttttatca gcaaagaata aaattaaatt aagaaggaca 2160  
 atggtgtccc aatccttata caaccaactt ccacaagaaa gtcaagtcag agacaacaaa 2220  
 aaaacaagca aaggaaattt ttaatttga gttgtcttgt ttgctgcata atttatgcag 2280

-continued

---

taaaacacta cacataaccc ttttagcagt agagcaatgg ttgaccgtgt gcttagcttc	2340
ttttatttta tttttttatc agcaaagaat aaataaaata aaatgagaca cttcagggat	2400
gtttcaacaa gcttggatcc tcgaagagaa gggttaataa cacacttttt taacattttt	2460
aacacaaatt ttagttatft aaaaatttat taaaaaattt aaaataagaa gaggaactct	2520
ttaaataaat ctaacttaca aaatttatga tttttaataa gttttcacca ataaaaaatg	2580
tcataaaaat atgttaaaaa gtatattatc aatattctct ttatgataaa taaaaagaaa	2640
aaaaaaataa aagttaagtg aaaatgagat tgaagtgact ttaggtgtgt ataaatata	2700
caacccgcc aacaatttat ttaatccaaa tatattgaag tatattatc catagccttt	2760
atftatttat atatttatta tataaaagct ttatttgttc taggtgttc atgaaatatt	2820
tttttggtt tatctccgtt gtaagaaaat catgtgcttt gtgtgccac tcaactattgc	2880
agctttttca tgcattggtc agattgacgg ttgattgtat ttttgtttt tatggttttg	2940
tgttatgact taagtcttca tctctttatc tcttcatcag gtttgatggt tacttaatat	3000
ggtccatggg tacatgcatg gttaaattag gtggccaact ttgtgtgtaa cgatagaatt	3060
ttttttatat taagtaaact atttttatat tatgaaataa taataaaaa aatattttat	3120
cattattaac aaaatcatat tagttaattt gtttaactta taataaaaga aatactgtaa	3180
cattcacatt acatggtaac atctttccac cctttcattt gtttttggt tgatgacttt	3240
ttttctggt taaatttatt tcccttcttt taaatttga atacattatc atcatatata	3300
aactaaaata ctaaaaacag gattacacaa atgataaata ataacacaaa tatttataaa	3360
tctagctgca atatatftaa actagctata tcgatattgt aaaataaac tagctgcaat	3420
gatactgata aaaaaatc atgtgctttc tggactgatg atgcagtata cttttgacat	3480
tgcttttatt ttatttttca gaaaagcttt cttagttctg ggttcttcat tatttgttc	3540
ccatctccat tgtgaattga atcatttgc tctgttcaca aatacaattt agntaggtac	3600
atgcattggt cagattcacg gtttattatg tcatgactta agttcatggt agtacattac	3660
ctgccacgca tgcattatat tggttagatt tgataggcaa atttggtgt caacaatata	3720
aatataaata atgtttttat attacgaat aacagtgatc aaaacaaaca gttttatctt	3780
tattaacaag attttgtttt tgtttgatga cgttttttaa tgtttacgct ttccccctc	3840
ttttgaattt agaacacttt atcatcataa aatcaaatc taaaaaatt acatatttca	3900
taaataataa cacaaatatt tttaaaaat ctgaaataat aatgaacaat attacatatt	3960
atcacgaaaa ttcattaata aaaatattat ataaataaaa tgtaatagta gttatatgta	4020
ggaaaaaagt actgcacgca taatatatac aaaaagatta aaatgaacta ttataaataa	4080
taacactaaa ttaatggtga atcatatcaa aataatgaaa aagtaataa aatttgtaat	4140
taacttctat atgtattaca cacacaaata ataaataata gtaaaaaaaa ttatgataaa	4200
tatttaccat ctcataagat atftaaaaa atgataaaaa tatagattat tttttatgca	4260
actagctagc caaaaagaga acacgggat atataaaaag agtaccttta aattctactg	4320
tacttcttt atctctgacg tttttatc aagtggacat acgtgaagat ttaattatc	4380
agtctaaata tttcattagc acttaatact tttctgtttt atctctatcc tataagtagt	4440
cccgattctc ccaacattgc ttattcacac aactaactaa gaaagtctc catagcccc	4500
caagcgccc atggcctct cggaggacgt catcaaggag tcatgcgct tcaagggtgcg	4560
catggagggc tccgtgaaag gccacgagtt cgagatcgag ggcgagggcg agggccgcc	4620
ctacgagggc acccagaccg ccaagctgaa ggtgaccaag ggcggcccc tgcccttgcg	4680

-continued

---

ctgggacatc	ctgtccccc	agttccagta	cggtccaag	gtgtacgtga	agcaccocgc	4740
cgacatcccc	gactacaaga	agctgtcctt	ccccgagggc	ttcaagtggg	agcgcgtgat	4800
gaacttcgag	gacggcggcg	tggtgaccgt	gaccocaggac	tcctocctgc	aggacggctc	4860
cttcatctac	aaggtgaagt	tcateggcgt	gaacttcccc	tcgacggcc	cgtaatgca	4920
gaagaagact	atgggctggg	aggcctccac	cgagcgcctg	tacccccgcg	acggcgtgct	4980
gaaggcgag	atccacaagg	ccctgaagct	gaaggacggc	ggccactacc	tggtggagtt	5040
caagtccatc	tacatggcca	agaagcccgt	gcagctgccc	ggctactact	acgtggactc	5100
caagctggac	atcacctccc	acaacgagga	ctacaccatc	gtggagcagt	acgagcgcgc	5160
cgagggccgc	caccacctgt	tcctgtagcg	gccggcccgc	acacaagtgt	gagagtacta	5220
aataaatgct	ttggtgtac	gaaatcatta	cactaaataa	aataatcaaa	gcttatatat	5280
gccttccgct	aaggccgaat	gcaaagaaat	tggttctttc	tcgttatctt	ttgccacttt	5340
tactagtacg	tattaattac	tacttaatca	tctttgttta	cggtcatta	tatccgtcga	5400
cggcgcgggc	cgctctagaa	ctagtggatc	cgtegacggc	gcgcccgate	atccggatat	5460
agttcctcct	ttcagcaaaa	aacctctcaa	gacctgttta	gaggcccaa	ggggttatgc	5520
tagttattgc	tcagcggtag	cagcagccaa	ctcagcttcc	tttcgggctt	tgtagcagc	5580
cggatcgatc	caagctgtac	ctcactatc	ctttgccctc	ggacgagtgc	tggggcgtcg	5640
gtttccacta	tcggcgagta	cttctacaca	gccatcggtc	cagacggccg	cgcttctgcg	5700
ggcgatttgt	gtacgccoga	cagtcccggc	tccggatcgg	acgattgcgt	cgcatcgacc	5760
ctgcccccaa	gctgcatcat	cgaattgcc	gtcaaccaag	ctctgataga	gttggtcaag	5820
accaatgcgg	agcatatacg	cccggagccg	cggcgatcct	gcaagctccg	gatgcctccg	5880
ctcgaagtag	cgcgtctgct	gctccataca	agccaaccac	ggcctccaga	agaagatggt	5940
ggcgacctcg	tattgggaat	ccccgaacat	cgctctgctc	cagtcaatga	ccgctgttat	6000
gcggccattg	tccgtcagga	cattgttggc	gccgaaatcc	gcgtgcacga	ggtgccggac	6060
ttcggggcag	tcctcggccc	aaagcatcag	ctcatcgaga	gcctgcgcga	cggacgcact	6120
gacggtgtcg	tccatcacag	tttgccagtg	atacacatgg	ggatcagcaa	tcgcgcatat	6180
gaaatcacgc	catgtagtgt	attgaccgat	tccttgcggt	ccgaatgggc	cgaaccgcgt	6240
cgctctggcta	agatcggccg	cagcagatgc	atccatagcc	tcgcgcaccg	gctgcagaac	6300
agcgggcagt	tcggtttcag	gcaggtcttg	caacgtgaca	ccctgtgcac	ggcgggagat	6360
gcaataggtc	aggctctcgc	tgaattcccc	aatgtcaagc	acttccggaa	tcgggagcgc	6420
ggccgatgca	aagtgccgat	aaacataacg	atctttgtag	aaaccatcgg	cgcagctatt	6480
taccgcgagg	acatatccac	gcctcctac	atcgaagctg	aaagcacgag	attcttcgcc	6540
ctccgagagc	tgcatcaggt	cggagacgct	gtcgaacttt	tcgatcagaa	acttctcgac	6600
agacgtcgcg	gtgagttcag	gcttttccat	gggtatatct	ccttcttaa	gttaaacaaa	6660
attatttteta	gagggaaacc	gttgtggtct	ccctatagtg	agtcgtatta	atttcgctgg	6720
atcgagatcg	atccaattcc	aatcccacaa	aaatctgagc	ttaacagcac	agttgctcct	6780
ctcagagcag	aatcgggtat	tcaacaccct	catatcaact	actacgttgt	gtataacggt	6840
ccacatgcgc	gtatatacga	tgactggggg	tgtacaaagg	cggcaacaaa	cggcgttccc	6900
ggagttgcac	acaagaaatt	tgccactatt	acagaggcaa	gagcagcagc	tgacgcgtac	6960
acaacaagtc	agcaaacaga	caggttgaac	ttcatcccca	aaggagaagc	tcaactcaag	7020

-continued

---

cccaagagct ttgctaagge cctaacaagc ccaccaaaagc aaaaagccca ctggctcacg	7080
ctaggaacca aaaggcccag cagtgatcca gcccctaaag agatctcctt tgcccggag	7140
attacaatgg acgatttctt ctatctttac gatctaggaa ggaagttcga aggtgaaggt	7200
gacgacacta tgttcaccac tgataatgag aaggtagcc tcttcaattt cagaaagaat	7260
gctgaccac agatggtag agaggcctac gcagcaggtc tcatcaagac gatctacccg	7320
agtaacaatc tccaggagat caaatcctt cccaagaagg ttaaagatgc agtcaaaaga	7380
ttcaggacta attgcatcaa gaacacagag aaagacatat ttctcaagat cagaagtact	7440
attccagtat ggacgattca aggcttgctt cataaaccaa ggcaagtaat agagattgga	7500
gtctctaaaa aggtagtctt tactgaatct aaggccatgc atggagtcta agattcaaat	7560
cgaggatcta acagaactcg ccgtgaagac tggcgaacag ttcatacaga gtcttttacg	7620
actcaatgac aagaagaaaa tcttcgtcaa catggtggag cacgacactc tggctactc	7680
caaaaatgac aaagatacag tctcagaaga ccaaagggtt attgagactt tcaacaaaag	7740
gataatttgc ggaacctcc tcggattcca ttgccagct atctgtcact tcatcgaag	7800
gacagtagaa aaggaaggtg gctcctacaa atgccatcat tgcgataaag gaaaggtat	7860
cattcaagat gcctctgccc acagtgttcc caaagatgga cccccacca cgaggagcat	7920
cgtggaaaaa gaagacgttc caaccacgtc ttcaaagcaa gtggattgat gtgacatctc	7980
cactgacgta agggatgacg cacaatccca ctatccttcg caagaccctt cctctatata	8040
aggaagtcca tttcatttgg agaggacacg ctcgagctca tttctctatt acttcagcca	8100
taacaaaaga actcttttct cttcttatta aaccatgaaa aagcctgaac tcaccgcgac	8160
gtctgtcgag aagtttctga tcgaaaagt cgcagcgcgc tccgacctga tgcagctctc	8220
ggagggcgaa gaatctctgt ctttcagctt cgatgtagga gggcgtggat atgtcctgcg	8280
ggtaaatagc tgcgccgatg gtttctacaa agatcgttat gtttatcggc actttgcatc	8340
ggccgcgctc ccgattccgg aagtgttga cattggggaa ttcagcgaga gcctgacctc	8400
ttgcatctcc cgcctgtcac aggggtgtcac gttgcaagac ctgcctgaaa ccgaactgcc	8460
cgctgttctg cagccggtcg cggaggccat ggatgcgacg gctgcggccc atcttagcca	8520
gacgagcggg ttcggcccat tcggaccgca aggaatcggc caatacacta catggcgtga	8580
tttcatatgc gcgattgtct atccccatgt gtatcactgg caaactgtga tggacgacac	8640
cgtcagtgcg tccgtcgcgc aggctctcga tgagctgatg ctttgggccc aggactgccc	8700
cgaagtccgg cacctcgtgc acgcggattt cggctccaac aatgtcctga cggacaatgg	8760
ccgcataaca gcggtcattg actggagcga ggcgatgttc ggggattccc aatcagaggt	8820
cgccaacatc ttcttctgga ggccgtggtt ggcttgtatg gagcagcaga cgcgctactt	8880
cgagcggagg catccggagc ttgcaggatc gccgcggctc cgggcgtata tgcctccgat	8940
tggctctgac caactctatc agagcttggc tgacggcaat ttcgatgatg cagcttgggc	9000
gcagggtcga tgcgacgcaa tcgtccgatc cggagccggg actgtcgggc gtacacaaat	9060
cgcccgcaga agcgcggccc tctggaccga tggctgtgta gaagtaactc ccgatagtgg	9120
aaaccgacgc cccagcactc gtccgagggc aaaggaatag tgaggtacct aaagaaggag	9180
tgcgtcgaag cagatcgttc aaacatttgg caataaagtt tcttaagatt gaatcctggt	9240
gccggtcttg cgatgattat catataattt ctgttgaatt acgttaagca tgtaataatt	9300
aacatgtaat gcatgacgtt atttatgaga tgggttttta tgattagagt cccgcaatta	9360
tacatttaat acgcgataga aaacaaaata tagcgcgcaa actaggataa attatcgcgc	9420

-continued

---

```

geggtgtcat ctatgttact agatcgatgt cgaatctgat caacctgcat taatgaatcg 9480
gccaacgcgc ggggagaggc ggtttgcgta ttgggcgctc ttccgcttcc tcgctcactg 9540
actcgctgcg ctcggtcggt cggctcgcgc gagcggatc agctcactca aaggcggtaa 9600
tacggttatc cacagaatca ggggataacg caggaaagaa catgtgagca aaaggccagc 9660
aaaagccag gaaccgtaaa aaggccgctg tgctggcggt tttccatagg ctccgcccc 9720
ctgacgagca tcacaaaaat cgacgctcaa gtcagagggt gcgaaaccgc acaggactat 9780
aaagatacca ggcggttccc cctggaagct cctcctgctg ctctcctggt ccgaccctgc 9840
cgcttaccgg atacctgtcc gcctttctcc ctccgggaag cgtggcgctt tctcaatgct 9900
cacgctgtag gtatctcagt tcgggtgtag tcgctcgctc caagctgggc tgtgtgcacg 9960
aacccccctg tcagcccgac cgctcgctc taccggtaa ctatcgtctt gagtccaacc 10020
cggaagaca cgacttatcg ccaactggcag cagccactgg taacaggatt agcagagcga 10080
ggtatgtagg cgggtgctaca gagttcttga agtggtggcc taactacggc tacactagaa 10140
ggacagtatt tggtatctgc gctctgctga agccagttac ctccgaaaaa agagttggta 10200
gctcttgatc cggcaaaaa accaccgctg gtagcgggtg ttttttgggt tgcaagcagc 10260
agattacgcg cagaaaaaaaa ggatctcaag aagatccttt gatcttttct acggggtctg 10320
acgctcagtg gaacgaaaa tcacgttaag ggattttggt catgacatta acctataaaa 10380
ataggcgtat cacgaggccc ttcgtctcg cgcgtttcgg tgatgacggt gaaaacctct 10440
gacacatgca gctcccggag acggctcacag cttgtctgta agcggatgcc gggagcagac 10500
aagcccgtca gggcgctca gcggtgttg gcggtgtcg gggctggctt aactatgctg 10560
catcagagca gattgtactg agagtgcacc atatggacat attgtcgta gaacgcggct 10620
acaattaata cataacctta tgtatcatac acatacgatt taggtgacac tatagaacgg 10680
cgcgccaagc ttttgatoca tgcccttcat ttgccccta ttaattaatt tggtaacagt 10740
ccgtactaat cagttactta tccttcccc atcataatta atcttggtag tctogaatgc 10800
cacaacactg actagtctct tggatcataa gaaaaagcca aggaacaaaa gaagacaaaa 10860
cacaatgaga gtatcctttg catagcaatg tctaagttca taaaattcaa acaaaaaacgc 10920
aatcacacac agtggacatc acttatccac tagctgatca ggatgcgcc gcgaagaaaa 10980
aaaaactgga ccccaaaagc catgcacaac aacacgtact cacaaagggt tcaatcgagc 11040
agcccaaac attcaccaac tcaaccctc atgagccctc acatttggtg tttctaacce 11100
aacctcaaac tcgtattctc ttccgccacc tcatttttgt ttatttcaac acccgtaaaa 11160
ctgcatgcca ccccgtyggc aaatgtccat gcatgttaac aagacctatg actataaata 11220
gctgcaatct cggcccaggt tttcatcctc aagaaccagt tcaatctct agtacaccgt 11280
attaagaat ttaagatata ctgc 11304

```

&lt;210&gt; SEQ ID NO 15

&lt;211&gt; LENGTH: 9147

&lt;212&gt; TYPE: DNA

&lt;213&gt; ORGANISM: Artificial Sequence

&lt;220&gt; FEATURE:

&lt;223&gt; OTHER INFORMATION: T-DNA of vector pZBL120xKS334

&lt;220&gt; FEATURE:

&lt;221&gt; NAME/KEY: misc\_feature

&lt;222&gt; LOCATION: (1892)..(1892)

&lt;223&gt; OTHER INFORMATION: n is a, c, g, or t

&lt;400&gt; SEQUENCE: 15

-continued

---

aattacaacg gatatatcc tgcgctcgac tctagaggat ccgcgccgtc gacggatcca	60
ctagtcttag agcgggccgc gccgctcgac gatataatga gccgtaaaca aagatgatta	120
agtagtaatt aatacgtact agtaaaagtg gcaaaagata acgagaaaga accaatttct	180
ttgcattcgg ccttagcggg aggcataat aagctttgat tattttattt agtgtaatga	240
tttcgtacaa ccaaagcatt tatttagtac tctcacactt gtgtcgccgc cggccgctac	300
aggaacaggt ggtggcggcc ctgcgcccgc tcgtactgct ccacgatggt gtagtctctg	360
ttgtgggagg tgatgtccag cttggagtcc acgtagtagt agccgggcag ctgcacgggc	420
ttcttgcca tgtagatgga cttgaactcc accaggtagt ggccgcccgc cttcagcttc	480
agggccttgt ggatctcgcc cttcagcagc ccgtcgccgg ggtacaggcg ctccgtggag	540
gcctcccagc ccatagtctt cttctgcatt acggggccgt cggaggggaa gttcacgccg	600
atgaacttca cctttagat gaaggagccg tctgcagggg aggagtctcg ggtcacggtc	660
accaagccgc cgtctcga gttcatcacg cgtcccact tgaagccctc ggggaaggac	720
agcttcttgt agtcggggat gtcggcgggg tgcttcacgt acaccttga gccgtaactg	780
aactgggggg acaggatgtc ccaggcgaag ggcagggggc cgcccttggc caccttcagc	840
ttggcggctt gggtgccctc gtagggggcg cctcgcctt cgccctcgat ctggaactcg	900
tggccgttca cggagccctc catgcgcacc ttgaagcga tgaactcctt gatgacgtcc	960
tcggaggagg ccatgggccc cttggggggc tatggaagac tttcttagtt agttgtgtga	1020
ataagcaatg ttgggagaat cgggactact tataggatag gaataaaaca gaaaagtatt	1080
aagtgctaat gaaatattta gactgataat taaaatcttc acgtatgtcc acttgatata	1140
aaaacgtcag gaataaagga agtacagtag aatttaaagg tactcttttt atatatacc	1200
gtgttctctt tttggctagc tagttgcata aaaaataatc tatattttta tcattatttt	1260
aaatatctta tgagatggta aatatttatc ataatttttt ttactattat ttattatttg	1320
tgtgtgtaat acatatagaa gtttaattaca aattttattt actttttcat tattttgata	1380
tgattcacca ttaatttagt gttattattt ataatagttc attttaatct ttttgatata	1440
attatgcgtg cagtactttt ttctacata taactactat tacattttat ttatataata	1500
tttttattaa tgaattttcg tgataaatg taatattggt cattattatt tcagattttt	1560
taaaaatatt tgtgttatta tttatgaaat atgtaatttt tttagtattt gattttatga	1620
tgataaagtg ttctaaatc aaaagaagg gaaagcgt aacattaaaa aacgtcatca	1680
aacaaaaaca aaactctgtt aataaagata aaactgtttg ttttgatcac tgttatttcg	1740
taatataaaa acattattta ttttatatt gttgacaacc aaatttgcc atcaaatcta	1800
accaatataa tgcatgcgtg gcaggtaatg tactaccatg aacttaagtc atgacataat	1860
aaaccgtgaa tctgaccaat gcattgtacct anctaaattg tatttgtgac acgaagcaaa	1920
tgattcaatt cacaatggag atgggaaaca aataatgaag aaccagaac taagaaagct	1980
tttctgaaaa ataaaaaaa ggcaatgtca aaagtatact gcacatcag tccagaaagc	2040
acatgatatt tttttatcag tatcaatgca gctagtttta ttttacaata tcgatatagc	2100
tagtttaaat atattgcagc tagatttata aatatttgtg ttattattta tcatttgtgt	2160
aatctgttt ttagtatttt agtttatata tgatgataat gtattccaaa tttaaaagaa	2220
gggaaataaa ttttaacaag aaaaaagtc atcaaacaaa aaacaaatga aagggtggaa	2280
agatgttacc atgtaatgtg aatgttacag tatttctttt attatagagt taacaaatta	2340
actaatatga ttttgtaat aatgataaaa tatttttttt attattattt cataatataa	2400

-continued

---

aaatagttta	cttaatatata	aaaaaattct	atcgttcaca	acaaagttgg	ccacctaat	2460
taaccatgca	tgtacctatg	gaccatatta	ggtaaccatc	aaacctgatg	aagagataaa	2520
gagatgaaga	cttaagtcac	aacacaaaac	cataaaaaac	aaaaatacaa	tcaaccgtca	2580
atctgaccaa	tgcataaaaa	agctgcaata	gtgagtgggc	acacaaagca	catgattttc	2640
ttacaacgga	gataaaaacca	aaaaaatatt	tcatgaacaa	cctagaacaa	ataaagcttt	2700
tatataataa	atatataaat	aaataaaggc	tatggaataa	tatacttcaa	tatatttgga	2760
ttaaataaat	tgttggcggg	gttgatata	ttatacacac	ctaaagtca	ttcaatctca	2820
ttttcactta	acttttattt	ttttttctt	tttatttctc	ataaagagaa	tattgataat	2880
atacttttta	acatattttt	atgacatttt	ttattggtga	aaacttatta	aaaatcataa	2940
atthtgaag	ttagattttt	ttaaagagtt	cctcttctta	ttttaaattt	tttaataaat	3000
ttttaaataa	ctaaaatttg	tgttaaaaat	gttaaaaaag	tgtgttatta	acccttctct	3060
tcgaggatcc	aagcttggg	aaacatccct	gaagtgtctc	atthtatttt	atthtattctt	3120
tgtctgataa	aaaataaaat	aaaagaagct	aagcacacgg	tcaaccattg	ctctactgct	3180
aaaagggtta	tgtgtagtgt	tttactgcat	aaattatgca	gcaacaaga	caactcaaat	3240
taaaaaattt	cctttgcttg	ttttttgtt	gtctctgact	tgactttctt	gtggaagttg	3300
gthgtataag	gattgggaca	ccattgtcct	tcttaattta	atthtattct	ttgctgataa	3360
aaaaaaaaat	ttcatatagt	gttaaataat	aatttgthta	ataacaaaa	agthcaaatat	3420
gtttactctc	gtttaaataa	ttgagattcg	tccagcaagg	ctaaacgatt	gtatagattt	3480
atgacaatat	ttactttttt	atagataaat	gttatattat	aataaattta	tatacatata	3540
ttatatgtta	tttattatta	ttttaaatcc	ttcaatattt	tatcaaacca	actcataatt	3600
ttttttttat	ctgtaagaag	caataaaatt	aaatagacct	actttaagga	tgatccaacc	3660
tttatacaga	gtaagagagt	tcaaatagta	ccctttcata	tacatatcaa	ctaaaatatt	3720
agaaatatca	tggatcaaac	cttataaaga	cattaaataa	gtggataagt	ataatatata	3780
aatgggtagt	atataatata	taaattggata	caaacttctc	tctttataat	tgttatgtct	3840
ccttaacatc	ctaatataat	acataagtg	gtaatatata	atatataaat	ggagacaaac	3900
ttcttccatt	ataattgtta	tgtcttctta	acacttatgt	ctcgttcaca	atgctaaggt	3960
tagaattgth	tagaaagtct	tatagtacac	atthgttttt	gtactatttg	aagcattcca	4020
taagccgtca	cgattcgat	gatttataat	aataagagga	aatttatcat	agaacaataa	4080
ggtgcataga	tagagtgtta	atatacata	acatcctttg	tttattcata	gaagaagtga	4140
gatggagctc	agttattata	ctgttacatg	gtcggataca	atattccatg	ctctccatga	4200
gctcttacac	ctacatgcat	tttagttcat	acttgccggc	gctagatcta	gagcatagtc	4260
acaagaaact	gtagttgtag	atgatgggtga	agaagcaaat	gatactatct	tctgcttgtt	4320
atccacatgc	atcatgcttc	caccagcccc	atthgtgtct	ccacactcaa	cactatccaa	4380
aactgcgctc	atattgttga	catcaactctc	agtctcatca	ccagggttgag	tgtggaacag	4440
aaaatcaatg	tcttctctca	acctgcaact	atcaaacatg	cccaccaagt	catcatcacc	4500
tttgagaaaa	gccaaagttg	gatcttgaaa	ctgagacaag	ccagtgatca	tgaagctcca	4560
tggcagatcc	tgtcaagca	catggtccat	aattgtgatc	agtgatgatg	attcttcttc	4620
ccttgagag	acatgggggg	tgtattgaac	ttggtgctgc	tgtggtggca	tgtgcagatt	4680
ttcagatggg	ggtggtgtgg	ttattgttgt	tgtctgtgtg	tctacttcta	cttcttcaga	4740

-continued

---

gtcagaggag ttaggaactg tttccgtttg tgcttctggt tgtgtgtgtt gggtttggtc	4800
atTTTTcttc tttattttgt ccatgtagtt gcttatgtca aaattggta ctgcattgac	4860
tccaagggtac tctattgctg ccatgtcata tgccactgct gcctcctctt gagttttata	4920
tgtcccaag tagaggtaact tgtttccgca tactcttcca attcgtgctt cccagcgacc	4980
attatgatga tgcctagcaa cccacaggta cttagacagg cctctagaaa agccactgct	5040
ttggcgccgc aaagaagcca aatattcttc tcttgaaacc ttgtocattt cctcgagctc	5100
cttggataaa gtttctatcg ggaaattcag ggttgcactt tttcccagc atttaagggc	5160
tgcaagggtca taggtacggg ctgcagattc ttcagatca tatgccccca aataaacttg	5220
tgcacccttc ttgctctgaa tgttgttcca agagetctta tcccataggt gagcttcaaa	5280
cctccctgtc cacctatgcc ttgtaactcc tctatagata gagcttcttc tgccaccagt	5340
ggtggtttgg ttctgcttgc atTTTTgtga cttcaaatta ttcctccttg gatgcttagg	5400
ccttcttttt tcaatggggag cttcaaacc aacagaggaa gtagatgatg aacaagaaga	5460
tgctggagac ctcttcatgc ggccgcagta tatcttaaat tctttaatac ggtgtagtag	5520
gatattgaac tggttcttga tgatgaaaac ctgggcccag attgcagcta tttatagtca	5580
taggtcttgt taacatgcat ggacatttgg ccacggggtg gcatgcagtt tgacgggtgt	5640
tgaaataaac aaaaatgagg tggcggaaga gaatacagat ttgaggttgg gttagaaca	5700
acaaatgtga gggctcatga tgggttgagt tggggaatgt tttgggctgc tcgattgaca	5760
cctttgtgag tacgtgttgt tgtgcatggt ttttggggtc cagttttttt ttcttgacgc	5820
ggcgatcctg atcagctagt ggataagtga tgtccactgt gtgtgattgc gtttttgttt	5880
gaattttatg aacttagaca ttgctatgca aaggatactc tcattgtggt ttgtcttctt	5940
ttgttccctg gctttttctt atgatccaag agactagtca gtgttgtggc attcgagact	6000
accaagatta attatgatgg ggaaggata agtaactgat tagtacggac tgttaccaaa	6060
ttaattaata agcggcaaat gaagggcatg gatcaaaagc ttggcgcgaa ttcactggcc	6120
gtcgttttac aacgtctgta ctgggaaaac cctggcggtta cccaacttaa tcgccttgca	6180
gcacatcccc ctttcgcccag ctggcgtaat agcgaagagg cccgcaccga tcgcccctcc	6240
caacagttgc gcagcctgaa tggcgaatgg atcgatccgt cgatcgacca aagcggccat	6300
cgtgcctccc cactcctgca gttcgggggc atggatgcgc ggatagccgc tgetggttct	6360
ctggatgccc acggatttgc actgccggtg gaactccgcg aggtcgtcca gcctcaggca	6420
gcagctgaac caactcgcga ggggatcgag cccctgctga gcctcgacat gttgtcgcaa	6480
aattgcctct ggacccgccc aacgatttgt cgtcactgtc aaggtttgac ctgcacttca	6540
tttggggccc acatacacca aaaaaatgct gcataattct cggggcagca agtcggttac	6600
ccggccgccc tgctggaccg ggttgaatgg tgcccgtaac tttcggtaga gcggacggcc	6660
aatactcaac ttcaaggaat ctaccccatg cgcgcggcgg gggaaaccgga gttcccttca	6720
gtgaacgtta ttagttcgcg gctcgggtgt tcgtagatac tagcccctgg ggccttttga	6780
aatttgaata agatttatgt aatcagtctt ttaggttga ccggttctgc cgcttttttt	6840
aaaattggat ttgtaataat aaaacgcaat tgtttgttat tgtggcgctc tatcatagat	6900
gtcgtataaa acctattcag cacaatatat tgttttcatt ttaatattgt acatataagt	6960
agtaggggtac aatcagtaaa ttgaacggag aatattatct ataaaaatac gatagtaacg	7020
ggtgatatat tcattagaat gaaccgaaac cggcggtaag gatctgagct acacatgctc	7080
aggtttttta caacgtgcac aacagaattg aaagcaata tcatgcatc ataggcgtct	7140

-continued

---

```

cgcatatctc attaagcag ggggtggcg aagaactcca gcatgagatc cccgcgctgg 7200
aggatcatcc agccggcgtc ccgaaaaacg attccgaagc ccaacctttc atagaaggcg 7260
gcggtggaat cgaaatctcg tgatggcagg ttgggcgctc cttggtcggt catttcgaac 7320
cccagagtcc cgctcagaag aactcgtcaa gaaggcgata gaaggcgatg cgctgcgaat 7380
cgggagcggc gataccgtaa agcacgagga agcggtcagc ccattcgccg ccaagctctt 7440
cagcaatata acgggtagcc aacgctatgt cctgatagcg gtccgccaca cccagccggc 7500
cacagtcatc gaatccagaa aagcggccat ttccaccat gatattcggc aagcaggcat 7560
cgccatgggt cacgacgaga tcctcgccgt cgggcacgcg cgccttgagc ctggcgaaca 7620
gttcggctgg cgcgagcccc tgatgctctt cgtccagatc atcctgatcg acaagaccgg 7680
cttccatccg agtacgtgct cgctcgatgc gatgtttcgc ttggtggtcg aatgggcagg 7740
tagccggatc aagcgtatgc agccgcgca ttgcatcagc catgatggat actttctcgg 7800
caggagcaag gtgagatgac aggagatcct gccccggcac ttcgccaat agcagccagt 7860
cccttcccgc ttcagtgaca acgtcgagca cagctgcgca aggaacgccc gtcgtggcca 7920
gccacgatag ccgcgctgcc tcgtcctgca gttcattcag ggcaccggac aggtcggctc 7980
tgacaaaaag aaccgggcgc cctcgcgctg acagccgaa cacggcggca tcagagcagc 8040
cgattgtctg ttgtgccag tcatagccga atagcctctc cacccaagcg gccggagaac 8100
ctgctgcaaa tccatcttgt tcaatcatgc gaaacgatcc ccgcaagctt ggagactggt 8160
gatttcagcg tgctctctcc aaatgaaatg aacttcctta tatagaggaa gggctcttgcg 8220
aaggatagtg ggattgtgcg tcaccctta cgtcagtgga gatatacat caatccactt 8280
gctttgaaga cgtggttggc acgtctctct tttccacgat gctcctcgtg ggtgggggtc 8340
catctttggg accactgtcg gcagagggcat cttcaacgat ggcccttccct ttatcgcaat 8400
gatggcattt gtaggagcca ccttcctttt ccaactatctt cacaataaag tgacagatag 8460
ctgggcaatg gaatccgagg aggtttccgg atattaccct ttgttgaaaa gtctcaattg 8520
ccctttggtc ttctgagact gtatctttga tatttttggc gtagacaagc gtgtcgtgct 8580
ccaccatggt gacgaagatt ttctctctgt cattgagtcg taagagactc tgtatgaact 8640
gttcgccagt ctttacggcg agttctgta ggtcctctat ttgaatcttt gactccatgg 8700
cctttgatcc agtgggaact acctttttag agactccaat ctctattact tgccctgggt 8760
tgtgaagcaa gccttgaatc gtcactactg gaatagtact tctgatcttg agaaatatat 8820
ctttctctgt gttcttgatg cagttagtcc tgaatctttt gactgcatct ttaaccttct 8880
tgggaaggta tttgatctcc tggagattat tgctcgggta gatcgtcttg atgagacctg 8940
ctgcgtaagc ctctctaacc atctgtgggt tagcattctt tctgaaattg aaaaggctaa 9000
tcttctcatt atcagtggtg aacatggat cgtcaccttc tccgtcgaac ttctgacta 9060
gatcgtagag atagaggaag tcgtccattg tgatctctgg ggcaaaggag atctgaatta 9120
tcatttacia ttgaatatat cctgcca 9147

```

&lt;210&gt; SEQ ID NO 16

&lt;211&gt; LENGTH: 3983

&lt;212&gt; TYPE: DNA

&lt;213&gt; ORGANISM: Artificial Sequence

&lt;220&gt; FEATURE:

&lt;223&gt; OTHER INFORMATION: vector pKR132

&lt;400&gt; SEQUENCE: 16

-continued

---

ctagagtcga cctgcaggca tgcaagcttg gcgtaatcat ggtcatagct gtttctctgtg	60
tgaaattggt atccgctcac aattccacac aacatacagag ccggaagcat aaagtgtaaa	120
gcctggggtg cctaatagagt gagctaactc acattaattg cgttgcgctc actgcccgct	180
ttccagtcgg gaaacctgtc gtgccagctg cattaatgaa tcggccaacg cgcggggaga	240
ggcggtttgc gtattggggc ctcttccgct tectcgtca ctgactcgtc gcgctcggtc	300
gttcggctgc ggcgagcgg atcagctcac tcaaaggcgg taatacgggt atccacagaa	360
tcaggggata acgcaggaaa gaacatgtga gcaaaaggcc agcaaaaggc caggaaccgt	420
aaaaaggcgg cgttgctggc gtttttccat aggctccgcc cccctgacga gcatcacaaa	480
aatcgacgct caagtcagag gtggcgaaaac ccgacaggac tataaagata ccaggcgttt	540
ccccctggaa gctccctcgt gcgctctcct gttccgacct tgccgcttac cggatactg	600
tccgccttcc tcccttgggg aagcgtggcg cttttctcata gctcacgctg taggtatctc	660
agttcgggtg aggtcgttcg ctccaagctg ggctgtgtgc acgaaacccc cgttcagccc	720
gaccgctcgg ccttatccgg taactatcgt cttgagttca acccgtaag acacgactta	780
tcgccactgg cagcagccac tggtaacagg attagcagag cgaggatgt aggcggtgct	840
acagagttct tgaagtggg gcctaactac ggctacacta gaaggacagt atttggatc	900
tgcgctctgc tgaagccagt tacctcggg aaaagagttg gtagctcttg atccggcaaa	960
caaaccaccg ctggtagcgg tggttttttt gtttgcaagc agcagattac gcgcagaaaa	1020
aaaggatctc aagaagatcc tttgatcttt tctacggggg ctgacgctca gtggaacgaa	1080
aactcacggt aagggatttt ggtcatgaga ttatcaaaaa ggatcttcac ctagatcctt	1140
ttaaataaaa aatgaagttt taaatcaatc taaagtatat atgagtaaac ttggtctgac	1200
agttaccaat gcttaatcag tgaggcacct atctcagcga tctgtctatt tcgttcatcc	1260
atagttgcct gactccccgt cgtgtagata actacgatac gggagggctt accatctggc	1320
cccagtgctg caatgatacc gcgagaccca cgctcaccgg ctccagattt atcagcaata	1380
aaccagccag ccggaagggc cgagcgcaga agtggtcctg caactttatc cgctccatc	1440
cagcttatta attgttgccg ggaagctaga gtaagtagtt cggcagttaa tagtttgccg	1500
aacgttgttg ccattgtctac aggcacgtg gtgtcaccgt cgctggttgg tatggttca	1560
ttcagctcgg gttcccaacg atcaaggcga gttacatgat ccccatggt gtgcaaaaaa	1620
gcggttagct ccttcggtcc tccgatcgtt gtcagaagta agttggccgc agtggtatca	1680
ctcatgggta tggcagcact gcataattct cttactgtca tgccatccgt aagatgcttt	1740
tctgtgactg gtgagtactc aaaccaagtca ttctgagaat agtgtatgcg gcgaccgagt	1800
tgctcttgcc cggcgtcaat acgggataat accgcgccac atagcagaac tttaaaagtg	1860
ctcatcattg gaaaacgttc ttcggggcga aaactctcaa ggatcttacc gctggtgaga	1920
tccagttcga tghtaaccac tcgtgcaccc aactgatctt cagcatcttt tactttcacc	1980
agcgtttctg ggtgagcaaa aacaggaagg caaaatgccg caaaaaaggg aataaggcgg	2040
acacggaat gttgaatact catactcttc ctttttcaat attattgaag catttatcag	2100
ggttattgtc tcatgagcgg atacatattt gaatgtattt agaaaaataa acaaataggg	2160
gttccgcgca catttccccg aaaagtgcc aetgacgtct aagaaacat tattatcatg	2220
acattaacct ataaaaatag gcgtatcacg aggcctcttc gtctcgcgcg tttcgggtgat	2280
gacggtgaaa acctctgaca catgcagctc ccggagacgg tcacagcttg tctgtaagcg	2340
gatgccggga gcagacaagc ccgtcagggc gcgtcagcgg gtgttgccgg gtgtcggggc	2400

-continued

---

tggetttaact atgcggcadc agagcagatt gtactgagag tgcacccat atgcgggtgtaa	2460
ataccgcaca gatgcgtaag gagaaaaac cgcacccaggc gccattcggc attcaggctg	2520
cgcaactggt ggggaaggcg atcgggtcgg gcctcttcgc tattacgcca gctggcgaaa	2580
gggggatgtg ctgcaaggcg attaagttgg gtaacgccag ggttttccca gtcacgacgt	2640
tgtaaaacga cggccagtga attcagagtc ggtaccggg gatcctctag acctgcaggc	2700
caactgcggt tggggctcca gattaaacga cgcctgttcg ttcctttcgc ttcacggctt	2760
aacgatgtcg tttctgtctg tgcccaaaaa ataaaggcat ttgttatttg caccagatat	2820
ttactaagtg caccctagtt tgacaagtag gcgataatta caaatagatg cggtgcaaat	2880
aataaatttt gaaggaaata attacaaaag aacagaactt atatttactt tattttaaaa	2940
aactaaaatg aaagaacaaa aaaagtaaaa aatacaaaaa atgtgcttta accactttca	3000
ttatttgta cagaaagtat gattctactc aaattgatct gttgtatctg gtgctgctt	3060
gtcacactgg cgatttcaat cccctaaaga tatggtgcaa actgcgaagt gatcaatatc	3120
tgctcggta atttagatta attaataata ttcaacgtga tgtacaaaa aaagacaatt	3180
ttttgctcca ttgacaaatt aaacctcctc aaggtaatth ccaaacctat aagcaaaaaa	3240
atctcacatt aattggcccg caatcctatt agtcttatta tactagagta ggaaaaaaa	3300
caattacaca acttgtctta ttattctcta tgctaatgaa tatttttccc ttttgtaga	3360
aatcagtggt tcctaattta ttgagtatta attcactca ccgcataat ttaccgttga	3420
ataagaaat tttacacata attcctttta agataaataa tttttttata ctgatctta	3480
tatgattacg tgaagccaag tgggttatac taatgatata taatgttga tagtaatcag	3540
tttataaacc aatgcatgg aatgttacg tggaaagcag taaattaaca agcattgaag	3600
caaatgcagc caccgcacca aaaccacccc acttcaactc cactgacat attocatgca	3660
actacaacac cctaaaaact caataaatgc ccccaacttc acttcaactc acccatcaat	3720
agcaagcggc cgcgaagtta aaagcaatgt tgtcacttgt cgtactaaca catgatgtga	3780
tagtttatgc tagctagcta taacataaagc tgtctctgag tgggttgatg attaataaag	3840
atcatcactg gtgaatggg atcgtgtacg taccctactt agtaggcaat ggaagcactt	3900
agagtgtgct ttgtgcatgg ccttgctctt gttttgagac ttttgaatg ttttcgagtt	3960
taaatctttg cctttgcgta cgt	3983

&lt;210&gt; SEQ ID NO 17

&lt;211&gt; LENGTH: 4746

&lt;212&gt; TYPE: DNA

&lt;213&gt; ORGANISM: Artificial Sequence

&lt;220&gt; FEATURE:

&lt;223&gt; OTHER INFORMATION: vector pKR627

&lt;400&gt; SEQUENCE: 17

gatcctctag acctgcaggc caactgcggt tggggctcca gattaaacga cgcctgttcg	60
ttcctttcgc ttcacggctt aacgatgtcg tttctgtctg tgcccaaaaa ataaaggcat	120
ttgttatttg caccagatat ttactaagtg caccctagtt tgacaagtag gcgataatta	180
caaatagatg cggtgcaaat aataaatttt gaaggaaata attacaaaag aacagaactt	240
atatttactt tattttaaaa aactaaaatg aaagaacaaa aaaagtaaaa aatacaaaaa	300
atgtgcttta accactttca ttatttgta cagaaagtat gattctactc aaattgatct	360
gttgtatctg gtgctgctt gtcacactgg cgatttcaat cccctaaaga tatggtgcaa	420

-continued

---

actgcgaagt gatcaatata tgctcgggta atttagatta attaataata ttcaactgta	480
tgtaccaaaa aaagacaatt ttttgcctca ttgacaaatt aaacctatc aaggtaattt	540
ccaaacctat aagcaaaaaa atttcacatt aattggcccg caatcctatt agtcttatta	600
tactagagta ggaaaaaaaa caattacaca acttgtctta ttattctcta tgctaataaa	660
tatttttccc ttttggtaga aatcagtggt tcttaattta ttgagtatta attccactca	720
cgcataat ttaccgttga ataagaaaat ttacacata attcttttta agataaataa	780
ttttttata ctgatctta tatgattacg tgaagccaag tgggttatac taatgatata	840
taatgtttga tagtaatcag tttataaacc aaatgcatgg aaatggtacg tgggaagcacg	900
taaattaaca agcattgaag caaatgcagc caccgcacca aaaccacccc acttcacttc	960
cacgtacat attccatgca actacaacac cctaaaactt caataaatgc cccacacttc	1020
acttcacttc acccatcaat agcaagcggc cgcgaagtta aaagcaatgt tgtaacttgt	1080
cgtactaaca catgatgtga tagtttatgc tagctagcta taacataagc tgtctctgag	1140
tgtgttgtat attaataaag atcatcactg gtgaatggtg atcgtgtacg taccctactt	1200
agtaggcaat ggaagcactt agagtgtgct ttgtgcatgg ccttgccctc gttttgagac	1260
ttttgtaatg ttttcgagtt taaatctttg cctttgcgta cgtctagagt cgagcatgca	1320
tctagagggc ccaattcggc ctatagttag tcgtattaca attcactggc cgtcgtttta	1380
caacgtcgtg actgggaaaa cctcggcgtt acccaactta atcgccttgc agcacatccc	1440
cctttcggca gctggcgtaa tagcgaagag gcccgcccg atcgcccttc ccaacagttg	1500
cgcagcctat acgtacggca gtttaaggtt tacacctata aaagagagag ccgttatcgt	1560
ctgtttgtgg atgtacagag tgatattatt gacacgccgg ggcgacggat ggtgatcccc	1620
ctggccagtg cacgtctgct gtcagataaa gtctccgctg aactttaccc ggtggtgcat	1680
atcggggatg aaagctggcg catgatgacc accgatatgg ccagtgtgccc ggtctccgtt	1740
atcggggaag aagtggctga tctcagccac cgcgaaaatg acatcaaaaa cgccattaac	1800
ctgatgttct ggggaatata aatgtcaggc atgagattat caaaaaggat cttcacctag	1860
atccttttca cgtagaaagc cagtccgag aaacgggtgct gaccccgat gaatgtcagc	1920
tactgggcta tctggacaag ggaaaaacgca agcgcacaaga gaaagcaggt agcttcagc	1980
gggcttacct ggcgatagct agactggcg gttttatgga cagcaagcga accggaattg	2040
ccagctgggg cgcctctggt taaggttggg aagccctgca aagtaaaactg gatggcttc	2100
tcgcccgaag ggatctgatg gcgcagggga tcaagctctg atcaagagac aggatgagga	2160
tcgtttcgca tgattgaaca agatggattg cacgcaggtt ctccggccgc ttgggtggag	2220
aggctattcg gctatgactg ggcacaacag acaatcggct gctctgatgc cgcctgttc	2280
eggctgtcag cgcagggggc cccggttctt tttgtcaaga ccgacctgtc cgggtccctg	2340
aatgaaactg aagacgagc agcgcgggta tcgtggctgg ccacgacggg cgttccttgc	2400
gcagctgtgc tcgacgttgt cactgaagcg ggaaggact ggctgctatt gggcgaagtg	2460
ccggggcagg atctcctgtc atctcacctt gctcctgccc agaaagtatc catcatggct	2520
gatgcaatgc ggcggctgca tacgcttgat ccggctacct gccattcga ccaccaagcg	2580
aaacatcgca tcgagcgagc acgtactcgg atggaagccg gtcttgcga tcaggatgat	2640
ctggacgaag agcatcaggg gctcgcgcca gccgaactgt tcgccaggct caaggcgagc	2700
atgccgaagc gcgaggatct cgtcgtgacc catggcgatg cctgcttgc gaatatcatg	2760
gtgaaaaatg gcccttttc tggattcact gactgtggcc ggctgggtgt ggcggaccgc	2820

-continued

---

tatcaggaca tagcgttggc taccegtgat attgctgaag agcttggcgg cgaatgggct	2880
gaccgcttcc tcgtgcttta cggtatcgcc gctcccgatt cgcagcgcac cgccttctat	2940
cgccttcttg acgagttcct ctgaattatt aacgcttaca atttctgat gcggtatfff	3000
ctccttacgc atctgtgctg tatttcacac cgcatacagg tggcactfff cggggaaatg	3060
tgcgcggaac ccctatttgt ttatfffctt aaatacattc aaatatgtat ccgctcatga	3120
gacaataacc ctgataaatg cttcaataat agcacgtgag gagggccacc atggccaagt	3180
tgaccagtgc cgttccgggt ctcaccgcgc gcgacgtcgc cggagcggtc gagttctgga	3240
ccgaccggct cgggttctcc cgggacttcg tggaggacga cttcgcgggt gtggtccggg	3300
acgacgtgac cctgttcatc agcgcgggtcc aggaccaggt ggtgccggac aacaccctgg	3360
cctgggtgtg ggtgcgcggc ctggacgagc tgtacgccga gtggtcggag gtcgtgtcca	3420
cgaacttccg ggacgcctcc gggccggcca tgaccgagat cggcgcgag cctggggggc	3480
gggagtctgc cctgcgcgac ccggccggca actgcgtgca cttcgtggcc gaggagcagg	3540
actgacacgt gctaaaaact catttttaat ttaaaaggat ctagggtgaag atcctfff	3600
ataatctcat gaccaaaac ccttaacgtg agtttctgtt ccaactgagc tcagacccc	3660
tagaaaagat caaaggatct tcttgagatc cttfffctt cgcgtaatc tgctgcttgc	3720
aaacaaaaaa accaccgcta ccagcgggtg tttgtttgcc ggatcaagag ctaccaactc	3780
tttttccgaa ggttaactggc ttcagcagag cgcagatacc aaatactgtc cttctagtgt	3840
agccgtagtt aggccaccac ttaagaact ctgtagcacc gcctacatac ctgcctctgc	3900
taatcctgtt accagtggct gctgccagtg gcgataagtc gtgtcttacc gggttggact	3960
caagacgata gttaccggat aaggcgcagc ggtcgggctg aacgggggt tcgtgcacac	4020
agcccagctt ggagcgaacg acctacaccg aactgagata cctacagcgt gagctatgag	4080
aaagcgcac ccttcccga gggagaaagg cggacaggta tccggtaagc ggcagggtcg	4140
gaacaggaga gcgcacgagg gagcttccag ggggaaacgc ctggtatctt tatagtcctg	4200
tcgggtttcg ccacctctga cttgagcgtc gatttttgtg atgctcgtca ggggggcgga	4260
gcctatggaa aaacgccagc aacgcggcct ttttaagggt cctgggcttt tgctggcctt	4320
ttgctcacat gttcttctc gcgttatccc ctgattctgt ggataaccgt attaccgctt	4380
ttgagtgagc tgataccgct cgcgcagacc gaacgaccga gcgcagcagc tcagtgagcg	4440
aggaagcggg agagcgcaca atacgcaaac cgcctctccc cgcgcgttgg ccgattcatt	4500
aatgcagctg gcacgacagg tttcccagct ggaagcggg cagtgagcgc aacgcaatta	4560
atgtgagtta gctcactcat taggcacccc aggttttaca ctttatgctt ccggctcgt	4620
tgttgtgtgg aattgtgagc ggataacaat ttcacacagg aaacagctat gaccatgatt	4680
acgccaagct atttaggtga cgcgtagaa tactcaagct atgcatcaag cttggtaccg	4740
agctcg	4746

&lt;210&gt; SEQ ID NO 18

&lt;211&gt; LENGTH: 4330

&lt;212&gt; TYPE: DNA

&lt;213&gt; ORGANISM: Artificial Sequence

&lt;220&gt; FEATURE:

&lt;223&gt; OTHER INFORMATION: vector KS294

&lt;400&gt; SEQUENCE: 18

agcttgaat tcgggatctg agtctagaaa tccgtcaaca tgggtggagca cgacactctc 60

-continued

---

gttactcca	agaatatcaa	agatacagtc	tcagaagacc	aaagggtat	tgagactttt	120
caacaaagg	taatategg	aaacctctc	ggattocatt	gcccagctat	ctgtcacttc	180
atcaaaagga	cagtagaaaa	ggaaggtg	acctacaaat	gccatcattg	cgataaagga	240
aaggctatcg	ttcaagatgc	ctctgcgac	agtggtecca	aagatggacc	cccaccacg	300
aggagcatcg	tggaaaaaga	agacgttcca	accacgtctt	caaagcaagt	ggattgatgt	360
gatgatccta	tgcgatgg	atgacgtgtg	ttcaagatga	tgacttcaaa	cctacctatg	420
acgtatggta	tgaacgtgtg	tcgactgatg	acttagatcc	actcgagcgg	ctataaatac	480
gtacctacgc	accctgcgct	accatcccta	gagctgcagc	ttatttttac	aacaattacc	540
aacaacaaca	aacaacaac	aacattacaa	ttactattta	caattacagt	cgaccggga	600
tcgtacctct	agggtggcgg	ccgcaagat	gaactaaaat	gcatgtaggt	gtaagagctc	660
atggagagca	tggaatattg	tatccgacca	tgtaacagta	taataactga	gctccatctc	720
actttcttcta	tgaataaaca	aaggatgta	tgatatatta	acactctatc	tatgcacctt	780
attgttctat	gataaatttc	ctcttattat	tataaatcat	ctgaatcgtg	acggcttatg	840
gaatgcttca	aatagtacaa	aaacaaatgt	gtactataag	actttctaaa	caattctaac	900
cttagcattg	tgaacgagac	ataagtgtta	agaagacata	acaattataa	tggagaaggt	960
ttgtctccat	ttatatatta	tatattacc	acttatgtat	tatattagga	tgtaaggag	1020
acataacaat	tataaagaga	gaagtttga	tccatttata	tattatatac	taccattta	1080
tatattatac	ttatccactt	athtaagtgc	tttataaggt	tgatccatg	atatttctaa	1140
tattttagtt	gatatgtata	tgaagggtta	ctatttgaac	tctcttactc	tgtataaagg	1200
ttggatcatc	cttaagtg	gtctattttaa	ttttattgct	tcttacagat	aaaaaaaaa	1260
ttatgagttg	gtttgataaa	atattgaagg	atthaaata	ataataaata	acataataa	1320
tatgtatata	aatttattat	aatataacat	ttatctataa	aaaagtaaat	attgtcataa	1380
atctatacaa	tcgtttagcc	ttgctggacg	aatctcaatt	atthaaacga	gagtaaacat	1440
atttgacttt	ttggttattt	aacaaattat	tatttaacac	tatatgaaat	ttttttttt	1500
atcagcaaag	aataaaatta	aattaagaag	gacaatggtg	tccaatcct	tatacaacca	1560
acttccacaa	gaaagtcaag	tcagagacaa	caaaaaaca	agcaaaggaa	atthtttaat	1620
ttgagttgtc	ttgtttgctg	cataatttat	gcagtaaac	actacacata	accttttag	1680
cagtagagca	atggttgacc	gtgtgcttag	cttcttttat	tttatthttt	tatcagcaaa	1740
gaataaataa	aataaaatga	gacacttcag	ggatgttca	acaagctcta	gactggaatt	1800
cgtagacggc	gccccgac	atccggatat	agttctctct	ttcagcaaaa	aacctctca	1860
gacctgttta	gagccccaa	gggttatgc	tagttattgc	tcagcgggtg	cagcagccaa	1920
ctcagcttcc	tttcgggctt	tgttagcagc	eggatcgatc	caagctgtac	ctcactattc	1980
ctttgccctc	ggacgagtg	tggggcgtcg	gtttccacta	tcggcgagta	cttctacaca	2040
gccatcggtc	cagacggcgg	cgcttctg	ggcgatttgt	gtacgccga	cagtcccggc	2100
tccggatcgg	acgattgctg	cgcatcgacc	ctgcgcccaa	gctgcatcat	cgaaattgcc	2160
gtcaaccaag	ctctgataga	gttggtcaag	accaatgcgg	agcatatc	cccggagccg	2220
egggatcct	gcaagctccg	gatgcctccg	ctcgaagtag	cgctctgct	gctccataca	2280
agccaaccac	ggcctccaga	agaagatgtt	ggcgacctcg	tattgggaat	ccccgaacat	2340
cgctcgtc	cagtcaatga	ccgctgttat	gcggccattg	tccgtcagga	cattgttgg	2400
gccgaaatcc	gcgtgcaoga	ggtgccggac	ttcggggcag	tcctcgcccc	aaagcatcag	2460

-continued

---

```

ctcatcgaga gcctgcgoga cggacgcact gacgggtgctg tccatcacag tttgccagtg 2520
atacacatgg ggatcagcaa tcgcgcatat gaaatcacgc catgtagtgt attgaccgat 2580
tccttgccgt ccgaatgggc cgaaccogct cgtctggcta agatcggccg cagcgatcgc 2640
atccatagcc tccgcaocg gctgcagAAC agcgggcagt tcggtttcag gcaggtcttg 2700
caacgtgaca ccctgtgcac ggcgggagat gcaataggtc aggtctctgc tgaattcccc 2760
aatgtcaagc acttccgaa tcggggagcgc ggccgatgca aagtgccgat aaacataacg 2820
atcctttagt aaaccatcg cgcagctatt taccgcagg acatatccac gccctctac 2880
atcgaagctg aaagcacgag attcttcgcc ctccgagagc tgcacaggt cggagacgct 2940
gtcgaacttt tcgatcagaa acttctcgac agacgtcgcg gtgagttcag gcttttccat 3000
gggtatatct ccttcttaaa gttaaacaaa attatttcta gagggaaacc gttgtggtct 3060
ccctatagtg agtcgtatta atttcgggg atcgagatct gatcaacctg cattaatgaa 3120
tcggccaacg cgcggggaga ggcggtttgc gtattgggag ctcttccgct tcctcgtca 3180
ctgactcgct gcgctcggtc gttcggctgc ggcgagcgg atcagctcac tcaaaggcgg 3240
taatacggtt atccacagaa tcaggggata acgcaggaaa gaacatgtga gcaaaaggcc 3300
agcaaaaggc caggaacogt aaaaaggcgg cgttgctggc gtttttccat aggtccgccc 3360
cccctgacga gcatcacaaa aatcgacgct caagtcagag gtggcgaaac ccgacaggac 3420
tataagata ccaggcgttt cccctggaa gctccctcgt gcgctctcct gttccgacct 3480
tgccgcttac cggatactg tcgccttctc tcccttcggg aagcgtggcg ctttctcaat 3540
gctcacgctg taggtatctc agttcggtgt aggtcgttcg ctccaagctg ggctgtgtgc 3600
acgaaccccc cgttcagccc gaccgctgcg ccttatccgg taactatcgt cttgagtcca 3660
acccggtaag acacgactta tcgccactgg cagcagccac tggtaacagg attagcagag 3720
cgaggtatgt aggcgggtct acagagttct tgaagtggtg gcctaactac ggctacacta 3780
gaaggacagt atttggatc tcgcctctgc tgaagccagt taccttcgga aaaagagttg 3840
gtagctcttg atccggcaaa caaaccccg ctggtagcgg tggttttttt gtttgcaagc 3900
agcagattac gcgcagaaaa aaaggatctc aagaagatcc tttgatcttt tctacggggt 3960
ctgacgctca gtggaacgaa aactcacggt aagggatttt ggtcatgaca ttaacctata 4020
aaaataggcg tatcacgagg cccttctgct tcgcgcgttt cggatgatgac ggtgaaaacc 4080
tctgacacat gcagctccc gagacgggtc agcttctct gtaagcggat gccgggagca 4140
gacaagcccg tcagggcgcg tcagcgggtg ttggcgggtg tcggggctgg cttactatg 4200
cggcatcaga gcagattgta ctgagagtgc accatagga catattgtcg ttagaacgcg 4260
gctacaatta atacataacc ttatgtatca tacacatagc atttaggtga cactatagaa 4320
cggcgcgcca 4330

```

&lt;210&gt; SEQ ID NO 19

&lt;211&gt; LENGTH: 5195

&lt;212&gt; TYPE: DNA

&lt;213&gt; ORGANISM: Artificial Sequence

&lt;220&gt; FEATURE:

&lt;223&gt; OTHER INFORMATION: vector pKR1142

&lt;400&gt; SEQUENCE: 19

```

ctagagggcc caattcggcc tatagtgagt cgtattacaa ttcactggcc gtcggtttac 60
aacgtcgtga ctgggaaaac cctggcggtta cccaacttaa tcgccttgca gcacatcccc 120

```

-continued

---

ctttcgcag	ctggcgtaat	agcgaagagg	cccgaccga	tcgcccctcc	caacagttgc	180
gcagcctata	cgtacggcag	tttaaggttt	acacctataa	aagagagagc	cgttatcgtc	240
tgtttggtga	tgtacagagt	gatattattg	acacgccggg	gcgacggatg	gtgatcccc	300
tggccagtgc	acgtctgctg	tcagataaag	tctcccgatg	actttaccgg	gtggtgcata	360
tcggggatga	aagctggcgc	atgatgacca	ccgatatggc	cagtgtgccg	gtctccgtta	420
tcggggaaga	agtggctgat	ctcagccacc	gcgaaaatga	catcaaaaac	gccattaacc	480
tgatgttctg	gggaatataa	atgtcaggca	tgagattatc	aaaaaggatc	ttcacctaga	540
tccttttcac	gtagaaagcc	agtcgcgaga	aacggctgctg	accccgatg	aatgtcagct	600
actgggctat	ctggacaagg	gaaaacgcaa	gcgcaaagag	aaagcaggta	gcttgccagtg	660
ggcttacatg	gcgatagcta	gactgggccc	ttttatggac	agcaagcgaa	ccggaattgc	720
cagctggggc	gccctctggt	aagggtggga	agccctgcaa	agtaaaactgg	atggctttct	780
cgccgcccaag	gatctgatgg	cgcaggggat	caagctctga	tcaagagaca	ggatgaggat	840
cgtttcgcat	gattgaacaa	gatggattgc	acgcaggttc	tccggccgct	tgggtggaga	900
ggctattcgg	ctatgactgg	gcacaacaga	caatcggctg	ctctgatgcc	gccgtgttcc	960
ggctgtcagc	gcagggggcgc	ccggttcttt	ttgtcaagac	cgacctgtcc	ggtgccctga	1020
atgaactgca	agacgaggca	gcgcggctat	cgtggctggc	cacgacgggc	gttccttgccg	1080
cagctgtgct	cgacgttgtc	actgaagcgg	gaagggactg	gctgctattg	ggcgaagtgc	1140
cggggcagga	tctcctgtca	tctcaccttg	ctcctgccga	gaaagtatcc	atcatggctg	1200
atgcaatgcg	gcggctgcat	acgcttgatc	cggtacctg	cccattcgac	caccaagcga	1260
aacatcgcat	cgagcgagca	cgtactcgga	tggaaagccg	tcttgctgat	caggatgatc	1320
tggacgaaga	gcatacgggg	ctcgcgccag	ccgaaactgt	cgccaggctc	aaggcgagca	1380
tgcccagcgg	cgaggatctc	gtcgtgacct	atggcgatgc	ctgcttgccc	aatatcatgg	1440
tggaaaatgg	ccgcttttct	ggattcatcg	actgtggccc	gctgggtgtg	gcggaaccgct	1500
atcaggacat	agcgttggtc	accctgata	ttgctgaaga	gcttggcggc	gaatgggctg	1560
accgcttctc	cgtgctttac	ggtatcgccc	ctcccgatc	gcagcgcac	gccttctatc	1620
gccttcttga	cgagttcttc	tgaattatta	acgcttacia	tttctgatg	cggtattttc	1680
tccttacgca	tctgtgccc	atttcacacc	gcatacaggt	ggcacttttc	ggggaaatgt	1740
gcgcggaacc	cctatttggt	tatttttcta	aatacattca	aatatgtatc	cgctcatgag	1800
acaataaacc	tgataaatgc	ttcaataata	gcacgtgagg	agggccacca	tggccaagt	1860
gaccagtgcc	gttccggctg	tcaccgcgcg	cgacgtgcc	ggagcggctg	agttctggac	1920
cgaccggctc	gggttctccc	gggacttcgt	ggagacgac	ttcgccggctg	tggtccggga	1980
cgacgtgacc	ctgttcatca	gcgcgggtcca	ggaccaggtg	gtgccggaca	acaccctggc	2040
ctgggtgtgg	gtgcgcggcc	tggacgagct	gtacgccgag	tggctggagg	tcgtgtccac	2100
gaacttccgg	gacgcctccg	ggccggccat	gaccgagatc	ggcgagcagc	cgtggggggc	2160
ggagttccgc	ctgcgcgacc	cgcccgcaa	ctgcgtgcac	ttcgtggccc	aggagcagga	2220
ctgacacgtg	ctaaaacttc	attttttaatt	taaaaggatc	taggtgaaga	tcctttttga	2280
taatctcatg	accaaaaatcc	cttaacgtga	gttttctgctc	cactgagcgt	cagaccccgt	2340
agaaaagatc	aaaggatcct	cttgagatcc	ttttttctcg	cgcgtaactc	gctgcttgca	2400
aacaaaaaaa	ccaccgctac	cagcgggtgt	ttgtttgccc	gatcaagagc	taccaactct	2460
ttttccgaag	gtaactggct	tcagcagagc	gcagatacca	aatactgtcc	ttctagtgtg	2520

-continued

---

gccgtagtta	ggccaccact	tcaagaactc	tgtagcaccc	cctacatacc	tcgctctgct	2580
aatcctgtta	ccagtggctg	ctgccagtgg	cgataagtcg	tgtcttaccg	ggttggactc	2640
aagacgatag	ttaccggata	aggcgcagcg	gtcgggctga	acggggggtt	cgtgcacaca	2700
gcccagcttg	gagcgaacga	cctacaccga	actgagatac	ctacagcgtg	agctatgaga	2760
aagcgccaag	cttcccgaag	ggagaaaggc	ggacaggtat	ccggtaaagc	gcagggtcgg	2820
aacaggagag	cgacagagg	agcttccagg	gggaaacgcc	tggtatcttt	atagtccctg	2880
cgggtttcgc	cacctctgac	ttgagcgtcg	atTTTTgtga	tgctcgtcag	gggggaggag	2940
cctatgaaa	aacgccagca	acgcggcctt	tttacggttc	ctgggctttt	gctggccttt	3000
tgctcacatg	ttctttctcg	cgttatcccc	tgattctgtg	gataaccgta	ttaccgcctt	3060
tgagtgagct	gataccgctc	gcccgcagcc	aacgaccgag	cgacagcag	cagtgcagca	3120
ggaagcggaa	gagcgcctaa	tacgcaaacc	gcctctcccc	gcgcggtggc	cgattcatta	3180
atgcagctgg	cacgacaggt	ttcccgactg	gaaagcgggc	agtgcagcga	acgcaattaa	3240
tgtgagttag	ctcactcatt	aggcacccca	ggctttacac	tttatgett	cggctcgtat	3300
gttgtgtgga	attgtgagcg	gataacaatt	tcacacagga	aacagctatg	accatgatta	3360
cgccaagcta	tttaggtgac	gcgttagaat	actcaagcta	tgcatcaagc	ttggtaccga	3420
gctcggatcc	tctagaaatc	cgtcaacatg	gtggagcagc	acactctcgt	ctactccaag	3480
aatatcaaag	atacagcttc	agaagaccaa	agggtatttg	agacttttca	acaaagggta	3540
atatcgggaa	acctcctcgg	attccattgc	ccagctatct	gtcacttcat	caaaaggaca	3600
gtagaaaagg	aaggtggcac	ctacaaatgc	catcattgcg	ataaaggaaa	ggctatcgtt	3660
caagatgcct	ctgccgacag	tggtcccaaa	gatggacccc	caccacagag	gagcatcgtg	3720
gaaaaagaag	acgttccaac	cacgtcttca	aagcaagtgg	attgatgtga	tgatccctat	3780
cgtatggtat	gacgtgtggt	caagatgatg	acttcaaaac	tacctatgac	gtatggtatg	3840
aacgtgtgtc	gactgatgac	ttagatccac	tcgagcggct	ataaatacgt	acctacgcac	3900
cctcgcgtac	catccctaga	gctgcagctt	atTTTTtaca	caattacca	caacaacaaa	3960
caacaacaaa	cattacaatt	actatttaca	attacagtcg	accggggatc	gtacctctag	4020
ggtggcggcc	gcaagtatga	actaaaatgc	atgtaggtgt	aagagctcat	ggagagcatg	4080
gaatattgta	tccgaccatg	taacagtata	ataactgagc	tccatctcac	ttcttctatg	4140
aataaacaaa	ggatgttatg	atatattaac	actctatcta	tgcaccttat	tgttctatga	4200
taaatttcct	cttattatta	taaatcatct	gaatcgtgac	ggcttatgga	atgcttcaaa	4260
tagtacaaaa	acaaatgtgt	actataagac	tttctaaaca	attctaacct	tagcattgtg	4320
aacgagacat	aagtgtaag	aagacataac	aattataatg	gaagaagttt	gtctccattt	4380
atatattata	tattaccac	ttatgtatta	tattaggatg	ttaaggagac	ataacaatta	4440
taaagagaga	agtttgtatc	catttatata	ttatatacta	cccatttata	tattatactt	4500
atccacttat	ttaatgtcct	tataagggtt	gatccatgat	atTTctaata	ttttagtgtg	4560
tatgtatatg	aaagggtact	atTTgaaact	tcttactctg	tataaagggt	ggatcatcct	4620
taaagtgggt	ctatttaatt	ttattgcttc	ttacagataa	aaaaaaaaat	atgagttggt	4680
ttgataaaat	attgaaggat	ttaaaataat	aataaataac	atataatata	tgtatataaa	4740
tttattataa	tataacattt	atctataaaa	aagtaaataat	tgtcataaat	ctatacaatc	4800
gtttagcctt	gctggacgaa	tctcaattat	ttaaacgaga	gtaaacatat	ttgacttttt	4860

-continued

---

```

ggttatttaa caaattatta ttaacacta tatgaaattt tttttttat cagcaaagaa 4920
taaaattaa ttaagaagga caatggtgtc ccaatcctta tacaaccaac ttccacaaga 4980
aagtcaagtc agagacaaca aaaaaacaag caaaggaat tttttaattt gagttgtctt 5040
gtttgctgca taatttatgc agtaaaacac tacacataac ccttttagca gtagagcaat 5100
ggttgaccgt gtgcttagct tcttttattt tattttttta tcagcaaaga ataaataaaa 5160
taaaatgaga cacttcaggg atgtttcaac aagct 5195

```

```

<210> SEQ ID NO 20
<211> LENGTH: 8314
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: vector pKR1141
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: (6481)..(6481)
<223> OTHER INFORMATION: n is a, c, g, or t

```

```

<400> SEQUENCE: 20

```

```

gatccgtcga cggcgcgccc gatcatccgg atatagttcc tcctttcagc aaaaaacccc 60
tcaagaccgg tttagaggcc ccaaggggtt atgctagtta ttgctcagcg gtggcagcag 120
ccaactcagc ttcctttcgg gctttgtag cagccggatc gatccaagct gtacctcact 180
attcctttgc cctcggacga gtgctggggc gtcggtttcc actatcggcg agtacttcta 240
cacagccatc ggtccagacg gccgcgcttc tgcgggggat ttgtgtacgc ccgacagtcc 300
cggctccgga tcggacgatt gcgtcgcate gacctgcgc ccaagctgca tcatcgaat 360
tgccgtcaac caagctctga tagagtgggt caagaccaat gcggagcata tacgcccgga 420
gccgcggcga tcctgcaagc tccggatgcc tccgctcgaa gtagcgcgctc tgetgetcca 480
tacaagccaa ccacggcctc cagaagaaga tgttggcgac ctcgatttgg gaatccccga 540
acatcgcctc gctccagtc atgaccgctg ttatcgggcc attgtccgctc aggacattgt 600
tggagccgaa atcccgctgc acgaggtgcc ggacttcggg gcagtcctcg gcccaaagca 660
tcagctcatc gagagcctgc gcgacggacg cactgacggt gtcgtccatc acagtttgcc 720
agtgatacac atggggatca gcaatcgcgc atatgaaatc acgccatgta gtgtattgac 780
cgattccttg cggtcggaat gggccgaacc cgctcgtctg gctaagatcg gccgcagcga 840
tcgcatccat agcctccgcy accggctgca gaacagcggg cagttcggtt tcaggcaggt 900
cttgcaactg gacaccctgt gcacggcggg agatgcaata ggtcaggctc tcgctgaatt 960
ccccaatgtc aagcacttcc ggaatcggga gcgcgggcga tgcaaaagtgc cgataaacat 1020
aacgatcttt gtagaaacca tcggcgcagc tatttaccgg caggacatat ccacgcctc 1080
ctacatcgaa gctgaaagca cgagattctt cgcctccga gagctgcatc aggtcggaga 1140
cgctgtcgaa cttttcgatc agaaaactct cgacagacgt cgcgggtgagt tcaggctttt 1200
ccatgggtat atctccttct taaagttaa caaaattatt tctagagggg aaccgttgtg 1260
gtctccctat agtgagtcgt attaatttcg cgggatcgag atcgatccaa ttccaatccc 1320
acaaaaatct gagcttaaca gcacagttgc tcctctcaga gcagaatcgg gtattcaaca 1380
ccctcatatc aactactacg ttgtgtataa cggtcacat gccgggtatat acgatgactg 1440
gggttgtaga aaggcggcaa caaacggcgt tcccggagtt gcacacaaga aatttgccac 1500
tattacagag gcaagagcag cagctgacgc gtacacaaca agtcagcaaa cagacaggtt 1560
gaacttcac cccaaaggag aagctcaact caagcccaag agctttgcta aggcctaac 1620

```

-continued

---

aagcccacca	aagcaaaaag	cccactggct	cacgctagga	accaaaaaggc	ccagcagtga	1680
tccagcccca	aaagagatct	cctttgcccc	ggagattaca	atggacgatt	tcctctatct	1740
ttacgatcta	ggaaggaagt	tcgaagggtga	agggtgacgac	actatgttca	ccactgataa	1800
tgagaaggtt	agcctcttca	atttcagaaa	gaatgctgac	ccacagatgg	ttagagaggc	1860
ctacgcagca	ggtctcatca	agacgatcta	cccagtaac	aatctccagg	agatcaaata	1920
ccttcccaag	aaggttaaag	atgcagtcaa	aagattcagg	actaattgca	tcaagaacac	1980
agagaaagac	atattttctca	agatcagaag	tactattcca	gtatggacga	ttcaaggctt	2040
gcttcataaa	ccaaggcaag	taatagagat	tggagtctct	aaaaaggtag	ttcctactga	2100
atctaaggcc	atgcatggag	tctaagattc	aaatcgagga	tctaacagaa	ctcgccgtga	2160
agactggcga	acagttcata	cagagtcttt	tacgactcaa	tgacaagaag	aaaaatcttcg	2220
tcaacatggt	ggagcacgac	actctggtct	actccaaaaa	tgtcaaagat	acagtctcag	2280
aagaccaaaag	ggctattgag	acttttcaac	aaaggataat	ttcgggaaac	ctcctcggat	2340
tccattgccc	agctatctgt	cacttcatcg	aaaggacagt	agaaaaaggaa	ggtggctcct	2400
acaaatgcca	tcattgcat	aaaggaaagg	ctatcattca	agatgcctct	gccgacagtg	2460
gtcccaaaga	tggaccccca	cccacgagga	gcacgtgga	aaaagaagac	gttccaacca	2520
cgtcttcaaa	gcaagtggat	tgatgtgaca	tctccactga	cgtaagggat	gacgcacaat	2580
cccactatcc	ttcgcaagac	ccttctctta	tataaggaag	ttcatttcat	ttggagagga	2640
cacgctcgag	ctcatttctc	tattacttca	gccataacaa	aagaactctt	ttctcttctt	2700
attaaacat	gaaaaagcct	gaactcaccg	cgacgtctgt	cgagaagttt	ctgatcgaaa	2760
agttcgacag	cgtctccgac	ctgatgcagc	tctcggaggg	cgaagaatct	cgtgctttca	2820
gcttcgatgt	aggagggcgt	ggatatgtcc	tgcgggtaaa	tagctgcgcc	gatggtttct	2880
acaaagatcg	ttatgtttat	cggcactttg	catcggccgc	gctcccgatt	ccggaagtgc	2940
ttgacattgg	ggaattcagc	gagagcctga	cctattgcat	ctcccgccgt	gcacagggtg	3000
tcacgttgca	agacctgcct	gaaaccgaac	tgcccgtctg	tctgcagccg	gtcgcggagg	3060
ccatggatgc	gatcgctgcg	gccgatctta	gccagacgag	cgggttcggc	ccattcggac	3120
cgcaaggaat	cggccaatac	actacatggc	gtgatttcat	atgcgcgatt	gctgatcccc	3180
atgtgatca	ctggcaaaact	gtgatggacg	acaccgtcag	tgcgtccgtc	gcgcaggctc	3240
tcgatgagct	gatgctttgg	gccgaggact	gccccgaagt	ccggcacctc	gtgcacgagg	3300
atttcggctc	caacaatgtc	ctgacggaca	atggccgcat	aacagcggtc	attgactgga	3360
gcgagggcat	gttcggggat	tcccataacg	aggtegccaa	catcttcttc	tggaggccgt	3420
ggttggcttg	tatggagcag	cagacgcgct	acttcgagcg	gaggcatccg	gagcttgca	3480
gatgcgccg	gctccggggc	tatatgctcc	gcattggtct	tgaccaactc	tatcagagct	3540
tggttgacgg	caatttcgat	gatgcagctt	gggcgcaggg	tcgatgcgac	gcaatcgtcc	3600
gatccggagc	cgggactgtc	gggcgtacac	aaatcgccc	cagaagcgcg	gccgtctgga	3660
ccgatggctg	tgtagaagta	ctcgccgata	gtggaaaccg	acgccccagc	actcgtccga	3720
gggcaaaagga	atagtgaggt	acctaagaa	ggagtgcgct	gaagcagatc	gttcaaacat	3780
ttggcaataa	agtttcttaa	gattgaatcc	tgttgccggt	cttgcgatga	ttatcatata	3840
atttctgttg	aattacgtta	agcatgtaat	aattaacatg	taatgcagta	cgttatttat	3900
gagatggggtt	tttatgatta	gagtcccgc	attatacatt	taatacgcga	tagaaaaaaa	3960

-continued

---

aatatagcgc gcaaaactagg ataaattatc gcgcgcgggtg tcatctatgt tactagatcg	4020
atgtcgaatc gatcaactcg cattaatgaa tcggccaacg cgcggggaga ggcggtttgc	4080
gtattggcgc ctcttccgct tctcgcctca ctgactcgtc gcgctcggtc gttcggctgc	4140
ggcgagcggc atcagctcac tcaaaggcgg taatacggtt atccacagaa tcaggggata	4200
acgcaggaagaa gaacatgtga gcaaaaaggcc agcaaaaaggc caggaaccgt aaaaaggccg	4260
cgttgctggc gtttttccat aggctccgcc cccctgacga gcatcacaaa aatcgacgct	4320
caagtccagag gtggcgaaac ccgacaggac tataaagata ccaggcgttt cccctggaa	4380
gctccctcgt gcgctctcct gttccgaccc tgccgcttac cggatacctg tccgccttc	4440
tccctcggg aagcgtggcg ctttctcaat gctcacgctg taggtatctc agttcgggtg	4500
aggtcgttgc ctccaagctg ggctgtgtgc acgaaccccc cgttcagccc gaccgctgcg	4560
ccttatccgg taactatcgt cttgagtcca acccggttaag acacgactta tcgccactgg	4620
cagcagccac tggtaacagg attagcagag cgaggatgtg aggcgggtgc acagagttct	4680
tgaagtggcg gcctaactac ggctacacta gaaggacagt atttggatc tgcgctctgc	4740
tgaagccagt tacctcggga aaaagagttg gtagctcttg atccggcaaa caaaccaccg	4800
ctggtagcgg tggttttttt gtttgcaagc agcagattac gcgcagaaaa aaaggatctc	4860
aagaagatcc tttgatcttt tctacggggt ctgacgctca gtggaacgaa aactcacgtt	4920
aagggatttt ggtcatgaca ttaacctata aaaataggcg tatcacgagg ccctttcgtc	4980
tcgcgcggtt cggtgatgac ggtgaaaacc tctgacacat gcagctcccg gagacggtea	5040
cagcttctct gtaagcggat gccgggagca gacaagcccg tcagggcgcg tcagcgggtg	5100
ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc	5160
accatattgga catattgtcg ttagaacgcg gctacaatta atacataacc ttatgtatca	5220
tacacatacg atttaggtga cactatagaa cggcgcgcca agcttggatc tctgcagga	5280
tctggccggc cggatctcgt acggatcctc gaagagaagg gttaataaca cactttttta	5340
acatttttaa cacaaatfff agttatftaa aaatttatta aaaaatttaa aataagaaga	5400
ggaactcttt aaataaatct aacttcaaaa atttatgatt ttaataagt tttcaccat	5460
aaaaaatgtc ataaaaatat gttaaaaagt atattatcaa tattctcttt atgataaata	5520
aaaagaaaaa aaaaaataaa gttaaagtga aatgagattg aagtgacttt aggtgtgtat	5580
aaatatatca accccgcaa caatttattt aatccaata tattgaagta tattattcca	5640
tagcctttat ttatttatat atttattata taaaagcttt atttgttcta ggttgttcat	5700
gaaatatttt tttggtttta tctccgttgt aagaaaaatca tgtgctttgt gtcgccactc	5760
actattgcag ctttttcatg cattggtcag attgacggtt gattgtattt ttgtttttta	5820
tggttttgtg ttatgactta agtcttcatc tctttatctc ttcacaggt ttgatggta	5880
cctaatatgg tccatgggta catgcatggt taaattaggg ggccaacttt gttgtgaacg	5940
atagaatttt ttttatatta agtaaaactat ttttatatta tgaaataata ataaaaaaaa	6000
tattttatca ttattaacaa aatcatatta gttaatgtgt taactctata ataaaagaaa	6060
tactgtaaca ttcacattac atggtaacat ctttccaccc tttcattgt tttttgttg	6120
atgacttttt ttcttgttta aatttatffc ccttctttta aatttggaa acattatcat	6180
catatataaa ctaaaactact aaaaacagga ttacacaaat gataaataat aacacaaata	6240
tttataaatc tagctgcaat atattttaa ac tagctatata gatattgtaa aataaaacta	6300
gctgcattga tactgataaa aaaatatcat gtgctttctg gactgatgat gcagtatact	6360

-continued

---

```

tttgacattg cctttatfff atttttcaga aaagctttct tagttctggg ttcttcatta 6420
tttgttccc atctccattg tgaattgaat catttgcttc gtgtcacaaa tacaatttag 6480
ntaggtacat gcattggcca gattcacggg ttattatgtc atgacttaag ttcattgtag 6540
tacattacct gccacgcatg catttatattg gtttagattg ataggcaaat ttgggtgtca 6600
acaatataaa tataaataat gtttttatat tacgaataaa cagtgatcaa aacaaacagt 6660
tttatcttta ttaacaagat ttgtttttg ttgatgacg ttttttaatg tttacgcttt 6720
ccccctctt ttgaatttag aacctttat catcataaaa tcaaaacta aaaaaattac 6780
atatttcata aataataaca caaatatfff taaaaaatct gaaataataa tgaacaatat 6840
tacatattat cacgaaaatt cattaataaa aatattatat aaataaatg taatagtagt 6900
tatatgtagg aaaaaagtac tgcacgcata atatatataa aaagattaaa atgaactatt 6960
ataaataata aactaaatt aatggatgaat catatcaaaa taatgaaaa gtaataaaaa 7020
tttgaatta acttctatat gtattacaca cacaaataat aaataatagt aaaaaaatt 7080
atgataaata tttaccatct cataagatat ttaaaataat gataaaaaa tagattattt 7140
tttatgcaac tagctagcca aaaagagaac acgggtatat ataaaaagag tacctttaa 7200
ttctactgta ctctctttat tctgacggt ttatatcaa gtggacatac gtgaagattt 7260
taattatcag tctaaatatt tcattagcac ttaatacttt tctgttttat tctatccta 7320
taagttagtc cgattctccc aacattgctt attcacacaa ctaactaaga aagtcttcca 7380
tagcccccca agcggcccat ggctctctcc gaggacgcca tcaaggagtt catgcgcttc 7440
aagggtgcga tggagggctc cgtgaaacgc cacgagttcg agatcgaggg cgagggcgag 7500
ggccgcccct acgagggcac ccagaccgcc aagctgaagg tgaccaaggg cgccccctg 7560
cccttcgctt gggacatcct gtccccccag ttcagtagc gctccaaggt gtacgtgaag 7620
caccgcccg acatccccga ctacaagaag ctgtccttcc cggagggctt caagtgggag 7680
cgcgtagtga acttcagga cggcggcgtg gtgaccgtga cccaggactc ctccctgcag 7740
gacggctcct tcactacaaa ggtgaagttc atcgcggtga acttcccctc cgacggcccc 7800
gtaatgcaga agaagactat gggctgggag gctccaccg agcgctgta cccccgcgac 7860
ggcgtgctga agggcgagat ccacaaggcc ctgaaactga aggacggcgg ccaactactg 7920
gtggagtcca agtccatccta catggccaag aagcccgtgc agctgcccgg ctactactac 7980
gtggactcca agctggacat cacctccac aacgaggact acaccatcgt ggagcagtac 8040
gagcgcgccc agggccgcca ccacctgttc ctgtagcggc cggccgcgac acaagtgtga 8100
gagtactaaa taaatgcttt ggtgttacga aatcattaca ctaaataaaa taatcaaagc 8160
ttatatatgc ctccgctaa ggccgaatgc aaagaaattg gttctttctc gttatctttt 8220
gccactttta ctagtacgta ttaattacta cttaatcacc ttgttttacg gctcattata 8280
tccgtcagcg gcgcccggcg ctctagaact agtg 8314

```

```

<210> SEQ ID NO 21
<211> LENGTH: 35
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: PCR primer SuSy-5

```

```

<400> SEQUENCE: 21

```

```

tccgtcaggt ctactcttta catgtcttt actcc

```

35

-continued

---

<210> SEQ ID NO 22  
 <211> LENGTH: 34  
 <212> TYPE: DNA  
 <213> ORGANISM: Artificial Sequence  
 <220> FEATURE:  
 <223> OTHER INFORMATION: PCR primer SuSy-3  
  
 <400> SEQUENCE: 22  
  
 agcggcgcgcg attttttctc agaggcaaaa acac 34

<210> SEQ ID NO 23  
 <211> LENGTH: 5531  
 <212> TYPE: DNA  
 <213> ORGANISM: Artificial Sequence  
 <220> FEATURE:  
 <223> OTHER INFORMATION: vector pLF122  
  
 <400> SEQUENCE: 23  
  
 cctgaattct gcagatatcc atcacactgg cggcgcgctcg agcatgcate tagagggccc 60  
 aattcgcctc atagtgagtc gtattacaat tcaactggccg tcgttttaca acgtcgtgac 120  
 tgggaaaacc ctggcgttac ccaacttaat cgccttgccg cacatcccc tttgcccagc 180  
 tggcgtaata gcgaagaggc cgcaccgat cgcccttccc aacagttgcg cagcctatac 240  
 gtacggcagt ttaaggttta cacctataaa agagagagcc gttatcgtct gtttgtggat 300  
 gtacagagtg atattattga cagccggggg cgacggatgg tgatccccct ggccagtgca 360  
 cgtctgctgt cagataaagt ctcccgtgaa ctttaccggg tggtgcatat cggggatgaa 420  
 agctggcgca tgatgaccac cgatatggcc agtgtgccgg tctccgttat cggggaagaa 480  
 gtggctgata tcagccaccg cgaaaatgac atcaaaaacg ccattaacct gatgttctgg 540  
 ggaatataaa tgtcaggcat gagattatca aaaaggatct tcacctagat ccttttcacg 600  
 tagaaagcca gtccgcagaa acggtgctga ccccggatga atgtcagcta ctgggctatc 660  
 tggacaaggg aaaacgcaag cgcaaagaga aagcaggtag cttgcagtgg gcttacatgg 720  
 cgatagctag actggggcgt tttatggaca gcaagcgaac cgggaattgc agctggggcg 780  
 ccctctggta aggttgggaa gccctgcaaa gtaaaactgga tggctttctc gccgccaagg 840  
 atctgatggc gcaggggatc aagctctgat caagagacag gatgaggatc gtttcgcatg 900  
 attgaacaag atggattgca cgcaggttct cgggccgctt gggtggagag gctattcggc 960  
 tatgactggg cacaacagac aatcggctgc tctgatgccg ccgtgttccg gctgtcagcg 1020  
 cagggggccc cggttctttt tgtcaagacc gacctgtccg gtgccctgaa tgaactgcaa 1080  
 gacgaggcag cgcggctatc gtggctggcc acgacggggc ttcttgcgc agctgtgctc 1140  
 gacgttgtca ctgaagcggg aagggaactgg ctgctattgg gcgaagtgcc ggggcaggat 1200  
 ctctgtcatc ctcacctgac tctgcccag aaagtatcca tcatggctga tgcaatgccc 1260  
 cggtgcata cgcttgatcc ggctacctgc ccattcgacc accaagcgaa acatcgcate 1320  
 gagcgagcac gtactcggat ggaagccggt cttgtcgate aggatgatct ggacgaagag 1380  
 catcaggggc tcgcgccagc cgaactgttc gccaggctca aggcgagcat gcccgacggc 1440  
 gaggatctcg tcgtgaccca tggcagatcc tgcttgccga atatcatggt ggaaaatggc 1500  
 cgctttctcg gattcatcga ctgtggcccg ctgggtgtgg cggaccgcta tcaggacata 1560  
 gcgttggtca cccgtgatat tgctgaagag cttggcggcg aatgggctga ccgcttctc 1620  
 gtgctttaag gtatcggcgc tcccattcgc cagcgcacgc ccttctatcg ccttcttgac 1680  
 gatttctct gaattattaa cgcttacaat tcctctgatc ggtattttct ccttacgcat 1740

-continued

---

ctgtgcggtg	tttcacacgc	catacaggtg	gcacttttcg	gggaaatgtg	cgcggaaccc	1800
ctatttggtt	atTTTTctaa	atacattcaa	atatgtatcc	gctcatgaga	caataaccct	1860
gataaatgct	tcaataatag	cacgtgagga	gggccaccat	ggccaagtgt	accagtgcgc	1920
ttccgggtgct	caccgcgcg	gacgtcgccg	gagcggtcga	gttctggacc	gaccggctcg	1980
ggttctcccg	ggacttcgtg	gaggacgact	tcgccgggtg	ggtcggggac	gacgtgaccc	2040
tgttcatcag	cgcggtccag	gaccaggtgg	tgccggacaa	cacctggcc	tgggtgtggg	2100
tgcgcgccct	ggacgagctg	tacgccgagt	ggtcggaggt	cggtgccacg	aacttccggg	2160
acgcctccgg	gccggccatg	accgagatcg	gcgagcagcc	gtgggggccc	gagttcgccc	2220
tgcgcgaccc	ggccggcaac	tcggtgcact	tcgtggccga	ggagcaggac	tgacacgtgc	2280
taaaacttca	tttttaattt	aaaagatct	aggtgaagat	cctttttgat	aatctcatga	2340
ccaaaatccc	ttaacgtgag	ttttcgttcc	actgagcgtc	agaccccgtg	gaaaagatca	2400
aaggatcttc	ttgagatcct	ttttttctgc	gcgtaactctg	ctgcttgcaa	acaaaaaac	2460
caccgctacc	agcgggtggt	tgtttgccgg	atcaagagct	accaactctt	tttccgaagg	2520
taactggctt	cagcagagcg	cagataccaa	atactgtcct	tctagtgtag	ccgtagttag	2580
gccaccactt	caagaactct	gtagcaccgc	ctacatacct	cgctctgcta	atcctgttac	2640
cagtggtctg	tgccagtggc	gataagtcgt	gtcttaccgg	gttggactca	agacgatagt	2700
taccggataa	ggcgcagcgg	tcgggctgaa	cggggggttc	gtgcacacag	cccagcttgg	2760
agcgaacgac	ctacaccgaa	ctgagatacc	tacagcgtga	gctatgagaa	agcggccacg	2820
ttcccgaagg	gagaaaggcg	gacaggtatc	cggttaagcgg	cagggtcggg	acaggagagc	2880
gcacgagggg	gcttccaggg	ggaaacgcct	ggtatcttta	tagtctctgc	gggtttcgcc	2940
acctctgact	tgagcgtcga	ttttgtgat	gctcgtcagg	ggggcggagc	ctatggaaaa	3000
acgccagcaa	cgcgcccttt	ttacggttcc	tgggcttttg	ctggcctttt	gctcacatgt	3060
tctttcctgc	gttatccctt	gattctgtgg	ataaccgtat	taccgccttt	gagtgagctg	3120
ataccgctcg	ccgcagccga	acgaccgagc	gcagcagatc	agtgagcag	gaagcggaa	3180
agcggccaat	acgcaaacgc	cctctccccg	cgcggtggcc	gattcattaa	tgcagctggc	3240
acgacaggtt	tcccgactgg	aaagcgggca	gtgagcggaa	cgcaattaat	gtgagttagc	3300
tcactcatta	ggcaccocag	gctttacact	ttatgcttcc	ggctcgtatg	ttgtgtggaa	3360
ttgtgagcgg	ataacaattt	cacacaggaa	acagctatga	ccatgattac	gccaaagctat	3420
ttaggtgacg	cgttagaata	ctcaagctat	gcatcaagct	tggtagcgag	ctcggatcca	3480
ctagtaacgg	ccgccagtg	gctggaattc	aggctctgca	ggctactctt	ttacatgttc	3540
tttactccgt	ctcaaaattt	cctttttttg	ttggctctct	ccgaacgagt	tggagaaatc	3600
gttaacccta	atcgaagatc	tagattcctc	tacatacgtt	tgatctctct	ctcagtatgg	3660
attacaaagc	gccaaaggaga	tactactcac	acggagttgt	tgcgagacag	caagatttcg	3720
caacagatat	agttacgaga	agaagacctt	atgtccctta	cgaccgtcca	aataagtttt	3780
caaggagtct	ggtttgagc	tcaaaagagt	acaaatcacc	cgagggcaat	aatatgcca	3840
ggaccaatga	tgtgtcaccg	aaaccaccag	ttttaggttt	ggcgaggaag	aatgctgctt	3900
gtgggccaat	gagatcttct	agtctcagaa	aatgggtatg	taagtattgg	aaagatggaa	3960
agtgcaagag	gggtgagcag	tgccagttct	tacactcttg	gtcttgtttc	cctggatttg	4020
ccatggtagc	ttctcttgaa	gggcacaata	aggaactaaa	ggggatcgct	ctccctgagg	4080

-continued

---

```

gttcagataa actcttttca gtcagtattg atggtacatt gcgagtttgg gactgcaatt 4140
ctggtcagtg tgtacattcc atcaaccttg acgcagaagc agggctctcta atcagtgaag 4200
gcccttgggt tttccttggc ttgccaaacg ctataaagcg ttttaacgtt caaaccagtc 4260
aagatttgca tcttcaagca gcaggggttg ttggtcaggt gaatgcaatg actattgcaa 4320
acggaatgct ttttgctgga acaagtcttg gtagtatctt agtctggaaa gctactacag 4380
actctgagtc tgatccattc aaatacttga catctcttga gggacatagt ggtgaagtca 4440
cttgttttgc tgttgagggt caaatgctat actctggttc tgtcgataaa acaatcaaga 4500
tgtgggatct caacacctg caatgtataa tgacctgaa gcaacatacc ggcaactgtca 4560
cttactctt atgttgggat aaatgttga tategtctt cttggatggg accataaaag 4620
tttgggctta ttctgaaaac ggaatcttga aagttgttca aactcgaga caagaacaga 4680
gtagtgttca tgctctttct ggtatgcatg atgcagaagc caaacggata atattctgct 4740
cttaccaaaa cggaaaccgtt ggcattttcg acctaccatc ttttcaagaa agaggaagga 4800
tgttctctac gcacacgatc gccacactca caattggtcc tcaaggattg ttattcagtg 4860
gagacgagag tggtaacttg cgtgtatgga ccttagctgc tggcaacaaa gtttagtctt 4920
ttcgactaaa gaattctgat ttaattttgt ggtttatatg ttgagttaac tgtaagaga 4980
gttttatttt gtaatagggt tatcagtcaa taaacaatct ttgtatcaac caaatgtaat 5040
ttttctcgtt aattcgattt cagagttttt actttaagat aaacaaactc tttcacacat 5100
catttaatga aagtggagaa gcttaaaaaa caaacaaaga aactgatcca tttttggcgg 5160
gtcttctct actcttattc atagtgtta acgaactata gcgtaaaatt cagagcaagc 5220
gatctccgat ttgaacgtgg ctatcaccgg aggccacca ctacgggcca tacgctctaa 5280
gtgaggatta aagtgtcttg gtgggtgacgt tgaagaaact cgcccatggt ttttgttatc 5340
tctgcagcca agtgtcgttc tttcttggc acttctcacc aagctacagt gaatttaaaa 5400
atggcgtctt tctttgatct cgtatacata agctggattg gtttcttaa caaatcctc 5460
tccttttggg tcttctgggt ttgccttgta agtgtttgtg tttttgcctc tgagaaaaaa 5520
tcgcgccgc t 5531

```

&lt;210&gt; SEQ ID NO 24

&lt;211&gt; LENGTH: 6644

&lt;212&gt; TYPE: DNA

&lt;213&gt; ORGANISM: Artificial Sequence

&lt;220&gt; FEATURE:

&lt;223&gt; OTHER INFORMATION: vector pKR1155

&lt;400&gt; SEQUENCE: 24

```

ggccgcaagt atgaactaaa atgcatgtag gtgtaagagc tcatggagag catggaatat 60
tgtatccgac catgtaacag tataataact gagctccatc tcaactcttc tatgaataaa 120
caaaggatgt tatgatatat taacctcta tctatgcacc ttattgttct atgataaatt 180
tcctcttatt attataaate atctgaatcg tgacggctta tggaatgctt caaatagtac 240
aaaaacaat gtgtactata agactttcta aacaattcta accttagcat tgtgaacgag 300
acataagtgt taagaagaca taacaattat aatggaagaa gtttgtctcc atttatatat 360
tatatattac ccacttatgt attatattag gatgttaagg agacataaca attataaaga 420
gagaagtttg tatecattta tatattatat actaccatt tatatattat acttatocac 480
ttatttaatg tctttataag gtttgatcca tgatattctt aatattttag ttgatatgta 540
tatgaaaggg tactatttga actctcttac tctgtataaa ggttggatca tccttaagtt 600

```

-continued

---

gggtctat	tttattg	cttcttacag	ataaaaaaaaa	aattatgagt	tggtttgata	660
aaatattgaa	ggatttaaaa	taataataaa	taacatataa	tatatgtata	taaatttatt	720
ataatataac	atztatctat	aaaaaagtaa	atattgtcat	aatctatac	aatcgtttag	780
ccttgctgga	cgaatctcaa	ttatttaaac	gagagtaaac	atatttgact	ttttggttat	840
ttaacaaatt	attatttaac	actatatgaa	atTTTTTTTT	ttatcagcaa	agaataaaat	900
taaattaaga	aggacaatgg	tgtoccaatc	cttatacaac	caacttccac	aagaaagtca	960
agtcagagac	aacaaaaaaaa	caagcaagg	aaatttttta	atTTGAGTTG	tcttgtttgc	1020
tgcataat	atgcagtaaa	acactacaca	taaccctttt	agcagtagag	caatggttga	1080
ccgtgtgctt	agcttctttt	atTTTATTTT	tttatcagca	aagaataaat	aaaataaaat	1140
gagacacttc	agggatgttt	caacaagctc	tagagggccc	aattcgccct	atagtgagtc	1200
gtattacaat	tactggccg	tcgTTTTACA	acgtcgtgac	tggaaaacc	ctggcgttac	1260
ccaacttaat	cgccttgag	cacatcccc	tttcgccagc	tgccgtaata	gcgaagaggc	1320
ccgcaccgat	cgccttccc	aacagttgcg	cagctatac	gtacggcagt	ttaaggttta	1380
cacctataaa	agagagagcc	gttatcgtct	gtttgtggat	gtacagagtg	atattattga	1440
cacgccgggg	cgaaggatgg	tgatccccct	ggccagtgca	cgctcgtctg	cagataaagt	1500
ctcccgtaa	ctttaccogg	tggtgcata	cggggatgaa	agctggcgca	tgatgaccac	1560
cgatatggcc	agtggtccgg	tctccgttat	cggggaagaa	gtggctgac	tcagccaccg	1620
cgaaaatgac	atcaaaaacg	ccattaacct	gatgttctgg	ggaatataaa	tgtcaggcat	1680
gagattatca	aaaaggatct	tcacctagat	ccttttcacg	tagaaagcca	gtccgcagaa	1740
acggtgctga	ccccggatga	atgtcagcta	ctgggctatc	tggacaaggg	aaaacgcaag	1800
cgcaaagaga	aagcaggtag	cttgacagtgg	gcttacatgg	cgatagctag	actgggcggt	1860
tttatggaca	gcaagcgaac	cggaattgcc	agctggggcg	ccctctggta	aggttgggaa	1920
gccctgcaaa	gtaaaactgga	tggttttctc	gccgccaagg	atctgatggc	gcaggggatc	1980
aagctctgat	caagagacag	gatgaggatc	gtttcgcagt	attgaacaag	atggattgca	2040
cgcaggttct	ccggccgctt	gggtggagag	gctattcggc	tatgactggg	cacaacagac	2100
aatcggctgc	tctgatgccg	ccgtgttccg	gctgtcagcg	caggggcgcc	cggttctttt	2160
tgtcaagacc	gacctgtccg	gtgccctgaa	tgaactgcaa	gacgaggcag	cgcggctatc	2220
gtggctggcc	acgacggggc	ttccttgccg	agctgtgctc	gacgttgta	ctgaagcggg	2280
aagggaactg	ctgctattgg	gcgaagtgcc	ggggcaggat	ctcctgtcat	ctcaccttgc	2340
tcctgccgag	aaagtatcca	tcattggctga	tgcaatcggg	cggtgcata	cgcttgatcc	2400
ggctacctgc	ccattcgacc	accaagcgaa	acatcgcac	gagcagacac	gtactcggat	2460
ggaagccggt	cttgtcgatc	aggatgatct	ggacgaagag	catcaggggc	tcgcccagc	2520
cgaactgttc	gccaggctca	aggcagcat	gcccgaacgc	gaggatctcg	tcgtgaccca	2580
tgccgatgcc	tgcttgccga	atatcatggt	ggaaaatggc	cgcttttctg	gattcatcga	2640
ctgtggccgg	ctgggtgtgg	cggaccgcta	tcaggacata	gcgttggtca	cccgtgatat	2700
tgctgaagag	cttggcggcg	aatgggctga	ccgcttctc	gtgctttacg	gtatcggcgc	2760
tcccgatctg	cagcgcacgc	ccttctatcg	ccttcttgac	gagttctct	gaattattaa	2820
cgcttacaat	ttcctgatgc	ggtatTTTCT	ccttacgcac	ctgtgcccga	tttcacaccg	2880
catacaggtg	gcacttttctg	gggaaatgtg	cgcggaaccc	ctatttgttt	atTTTCTAA	2940

-continued

---

atacattcaa	atatgtatcc	gctcatgaga	caataaccc	gataaatgct	tcaataatag	3000
cacgtgagga	gggccaacat	ggccaagttg	accagtgccg	ttccgggtgct	caccgcgcgc	3060
gacgtcgcgc	gagcggtcga	gttctggacc	gaccggctcg	ggttctcccg	ggacttcgtg	3120
gaggacgact	tcgccggtgt	ggccggggac	gacgtgaccc	tgttcatcag	cgcggtccag	3180
gaccaggtgg	tgccggacaa	caccctggcc	tgggtgtggg	tgccggcct	ggacgagctg	3240
tacgccgagt	ggteggaggt	cgtgtccacg	aacttccggg	acgcctccgg	gcccggccatg	3300
accgagatcg	gagagcagcc	gtggggggcg	gagttcgccc	tgccggaccc	ggccggcaac	3360
tgcggtgact	tcgtggccga	ggagcaggac	tgacacgtgc	taaaactca	tttttaattt	3420
aaaaggatct	aggtaagat	cctttttgat	aatctcatga	ccaaaatccc	ttaacgtgag	3480
ttttcgttcc	actgagcgtc	agaccccgta	gaaaagatca	aaggatcttc	ttgagatcct	3540
ttttttctgc	gcgtaactcg	ctgcttgcaa	acaaaaaaaa	caccgctacc	agcggtggtt	3600
tgtttgccgg	atcaagagct	accaactctt	tttccgaagg	taactggctt	cagcagagcg	3660
cagataccaa	atactgtcct	tctagtgtag	ccgtagttag	gccaccactt	caagaactct	3720
gtagcaccgc	ctacatacct	cgctctgcta	atcctgttac	cagtggctgc	tgccagtggc	3780
gataagtcgt	gtcttaccgg	gttgactca	agacgatagt	taccggataa	ggcgcagcgg	3840
tcgggctgaa	cgggggggtt	gtgcacacag	cccagcttgg	agcgaacgac	ctacaccgaa	3900
ctgagatacc	tacagcgtga	gctatgagaa	agcgcaccgc	ttcccgaagg	gagaaaaggcg	3960
gacaggtatc	cggtaagcgg	cagggtcggg	acaggagagc	gcacgagggg	gcttccaggg	4020
ggaaaacgct	ggtatcttta	tagtctgtc	gggtttcgcc	acctctgact	tgagcgtcga	4080
tttttgtgat	gctcgtcagg	ggggcggagc	ctatggaaaa	acgccagcaa	cgcggccttt	4140
ttacggttcc	tgggcttttg	ctggcctttt	gctcacatgt	tctttcctgc	gttatcccct	4200
gattctgtgg	ataaccgtat	taccgccttt	gagtgagctg	ataccgctcg	ccgcagccga	4260
acgaccgagc	gcagcagatc	agtgcagcag	gaagcgggaag	agcgcaccaat	acgcaaacccg	4320
cctctccccg	cgcgttggcc	gattcattaa	tgacagctggc	acgacaggtt	tcccgactgg	4380
aaagcgggca	gtgagcgcga	cgcaattaat	gtgagttagc	tactcatta	ggcaccocag	4440
gctttacact	ttatgcttcc	ggctcgtatg	ttgtgtggaa	ttgtgagcgg	ataacaattt	4500
cacacaggaa	acagctatga	ccatgattac	gccaagctat	ttaggtgacg	cgttagaata	4560
ctcaagctat	gcatcaagct	tggtaccgag	ctcgatcca	ctagtaacgg	ccgcagtggt	4620
gctggaattc	aggctcctga	ggtctactct	ttacatgttc	tttactccgt	ctcaaaattt	4680
cctttttttg	ttggctctct	ccgaacgagt	tggagaaatc	gttaacccta	atcgaagatc	4740
tagattcctc	tacatacgtt	tgatctctct	ctcagtatgg	attacaaagc	gccaaggaga	4800
tactactcac	acggagttgt	tgccgagacag	caagatttcc	caacagatat	agttacgaga	4860
agaagacctt	atgtccctta	cgaccgtcca	aataagtttt	caaggagtct	ggtttgagcg	4920
tcaaaagagt	acaaatcacc	cgagggcaat	aatatgcca	ggaccaatga	tgtgtcaccg	4980
aaaccaccag	ttttaggttt	ggcagaggaag	aatgctgctt	gtgggccaat	gagatcttct	5040
agtctcagaa	aatgggtatg	taagtattgg	aaagatggaa	agtgcaagag	gggtgagcag	5100
tgccagttct	tacactcttg	gtcttgtttc	cctggattgg	ccatggtagc	ttctcttgaa	5160
gggcacaata	aggaactaaa	ggggatcgct	ctccctgagg	gttcagataa	actcttttca	5220
gtcagtatgg	atggtacatt	gagagtttgg	gactgcaatt	ctggtcagtg	tgtacattcc	5280
atcaaccttg	acgcagaagc	agggtctcta	atcagtgaag	gcccttgggt	ttccttggc	5340

-continued

---

```

ttgccaaaagc ctataaaggc ttttaacggt caaaccagtc aagatttgca ttttcaagca 5400
gcaggggtgg ttggtcaggt gaatgcaatg actattgcaa acggaatgct ttttgetgga 5460
acaagtctctg gtagtatctt agtctggaaa gctactacag actctgagtc tgatccattc 5520
aaatacttga catctcttga gggacatagt ggtgaagtca cttgttttgc tgttgagggt 5580
caaatgctat actctggttc tgcgataaa acaatcaaga tgtgggatct caacacctg 5640
caatgtataa tgaccctgaa gcaacatacc ggcactgtca cttcactctt atgttgggat 5700
aaatgtttga tategtcttc cttggatggg accataaaag tttgggctta ttctgaaaac 5760
ggaatcttga aagttgttca aactcgcaga caagaacaga gtagtggtca tgctctttct 5820
ggtatgcatg atgcagaagc caaaccgata atattctgct cttacaaaaa cggaaccgtt 5880
ggcattttcg acctaccatc ttttcaagaa agaggaagga tgttctctac gcacacgatc 5940
gccacactca caattggtcc tcaaggattg ttattcagtg gagacgagag tggttaactg 6000
cgtgtatgga ccttagctgc tggcaacaaa gtttagtctt ttcgactaaa gaattctgat 6060
ttaattttgt ggtttatgat ttgagttaac tgtaagaga gttttatctt gtaatagggtg 6120
tatcagtc aaacaatct ttgtatcaac caaatgtaat ttttctcggt aattcgattt 6180
cagagttttt actttaagat aaacaaactc tttcacacat catttaatga aagtggagaa 6240
gcttaaaaaa caaacaaga aactgatcca tttttggcgg gtctctctct actcttattc 6300
atatgtggtta acgaactata gcgtaaaatt cagagcaagc gatctccgat ttgaacgtgg 6360
ctatcaccgg aggccacca ctacgggcca tacgctctaa gtgaggatta aagtgetctg 6420
gtggtgacgt tgaagaaact cgcctatggt ttttgttacc tctgcagcca agtgcgttc 6480
tttcttcgcc acttctcacc aagctacagt gaatttaaaa atggcgtctt tctttgatct 6540
cgtatacata agctggattg gtttcttaaa caaatcctc tccttttggg tcttctgggt 6600
tgcccttgta agtgtttgtg tttttgctc tgagaaaaaa tcgc 6644

```

```

<210> SEQ ID NO 25
<211> LENGTH: 11736
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: vector pKR1158
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: (7136)..(7136)
<223> OTHER INFORMATION: n is a, c, g, or t

<400> SEQUENCE: 25

```

```

gtacgagatc cggccggcca gatcctgcag gagatccaag cttggcgcgc cgttctatag 60
tgtcacctaa atcgatgtg tatgatacat aaggttatgt attaattgta gccgcgttct 120
aacgacaata tgtccatgat gtgcactctc agtacaatct gctctgatgc cgcatagtta 180
agccagcccc gacacccgcc aacaccgct gacgcgccct gacgggcttg tctgctcccc 240
gcatccgctt acagacaagc tgtgacgctc tccgggagct gcatgtgtca gaggttttca 300
ccgtcatcac cgaaacgcgc gagacgaaag ggcctcgtga tacgcctatt tttatagggt 360
aatgtcatga ccaaaatccc ttaacgtgag ttttctgtcc actgagcgtc agaccccgta 420
gaaaagatca aaggatcttc ttgagatcct tttttctgc gcgtaatctg ctgcttgcaa 480
acaaaaaac caccgctacc agcggtggtt tgtttgccgg atcaagagct accaactctt 540
tttccgaagg taactggctt cagcagagcg cagataccaa atactgtcct tctagtgtag 600

```

-continued

---

ccgtagttag gccaccactt caagaactct gtagcaccgc ctacatacct cgctctgcta	660
atcctgttac cagtggctgc tgccagtggc gataagtcgt gtcttacccg gttggactca	720
agacgatagt tacccgataa ggcgcagcgg tcgggctgaa cggggggttc gtgcacacag	780
cccagcttgg agcgaacgac ctacaccgaa ctgagatacc tacagcgtga gcattgagaa	840
agcgcaccgc tccccgaagg gagaaaggcg gacaggtatc cggtaagcgg cagggtcgga	900
acagagagac gcacgagggg gcttccaggg ggaacgcct ggatcttta tagtcctgtc	960
gggttccgcc acctctgact tgagcgtcga tttttgtgat gctcgtcagg ggggcccggc	1020
ctatggaaaa acgccagcaa cgcggccttt ttacggttcc tggccttttg ctggcctttt	1080
gctcacatgt tcttctctgc gttatcccct gattctgtgg ataaccgtat tacgccttt	1140
gagtgagctg ataccgctcg ccgcagccga acgaccgagc gcagcagtc agtgagcag	1200
gaagcgggaag agcgcaccaat acgcaaaccc cctctccccg cgcgttggcc gattcattaa	1260
tcagggttga tcgattcgc acgatctag taacatagat gacaccgcgc gcgataattt	1320
atcctagttt gcgcgctata ttttgtttc tatcgcgtat taaatgtata attgcgggac	1380
tctaatacata aaaaccatc tcataaataa cgtcatgcat tacatgttaa ttattacatg	1440
cttaacgtaa ttcaacagaa attatatgat aatcatcgca agaccggcaa caggattcaa	1500
tcttaagaaa ctttattgac aaatgttga acgatctgct tcgacgcact ccttctttag	1560
gtacctcact attcctttgc cctcggacga gtgctggggc gtcggtttcc actatcggcg	1620
agtacttcta cacagccatc ggtccagacg gccgcgcttc tgccggcgat ttgtgtacgc	1680
ccgacagtcc cggctccgga tcggacgatt gcgtcgcac gaccctgcgc ccaagctgca	1740
tcacgaaat tcgcgtcaac caagctctga tagagttggt caagaccaat gcggagcata	1800
tacgcccgga gcccgggcga tctgcaagc tccggatgcc tccgctcgaa gtagecgcgc	1860
tgctgctcca tacaagccaa ccacggcctc cagaagaaga tgttggcgac ctcgattgg	1920
gaatccccga acatcgctc gctccagtc atgaccgctg ttatgcggcc attgtccgctc	1980
aggacattgt tggagccgaa atccgcgtgc acgaggtgcc ggacttcggg gcagtcctcg	2040
gccccaaagca tcagctcatc gagagcctgc gcgacggacg cactgacggg gtcgtccatc	2100
acagtttgcc agtgatacac atggggatca gcaatcgcgc atatgaaatc acgccatgta	2160
gtgtattgac cgattccttg cggtcgcaat gggccgaacc cgcctgctcg gctaagatcg	2220
gccgcagcga tcgcatccat ggcctccgcg accggtgca gaacagcggg cagttcgggt	2280
tcaggcaggt cttgcaacgt gacaccctgt gcacggcggg agatgcaata ggtcaggctc	2340
tcgctgaatt ccccaatgtc aagcaattcc ggaatcggga gcgcggccga tgcaaatgtc	2400
cgataaacat aacgatcttt gtagaacca tcggcgcagc tatttaccgc caggacatat	2460
ccacgccctc ctacatcgaa gctgaaagca cgagattctt cgcctccga gagctgcatc	2520
aggtcggaga cgcgtcgcga cttttcgcac agaaacttct cgacagacgt cgcggtgagt	2580
tcaggctttt tcatggttta ataagaagag aaaagagttc ttttgttatg gctgaagtaa	2640
tagagaaatg agctcgcagc gtcctctcc aatgaaatg aacttcotta tatagaggaa	2700
gggtcttgcg aaggatagtg ggattgtgcg tcatccctta cgtcagtgga gatgtcacat	2760
caatccactt gctttgaaga cgtggttga acgtctctt tttccacgat gctcctcgtg	2820
gggtggggtc catctttggg accactgtcg gcagaggcat cttgaatgat agcctttcct	2880
ttatcgcaat gatggcattt gtaggagcca ccttcctttt ctactgtcct ttcgatgaag	2940
tgacagatag ctgggcaatg gaatccgagg aggtttcccg aaattatcct ttgttggaaa	3000

-continued

---

gtctcaatag ccctttggtc ttctgagact gtatctttga catttttggg gtagaccaga	3060
gtgtcgtgct ccaccatggt gacgaagatt ttcttcttgt cattgagtcg taaaagactc	3120
tgtatgaact gttcgccagt ctccacggcg agttctgtta gatcctcgat ttgaatctta	3180
gactccatgc atggccttag attcagtagg aactaccttt ttagagactc caatctctat	3240
tacttgccctt ggtttatgaa gcaagccttg aatcgccat actggaatag tacttctgat	3300
cttgagaaat atgtctttct ctgtgttctt gatgcaatta gtccctgaatc ttttgactgc	3360
atctttaacc ttcttgggaa ggtatttgat ctccctggaga ttgttactcg ggtagatcgt	3420
cttgatgaga cctgctgogt aggcctctct aaccatctgt gggtcagcat tctttctgaa	3480
attgaagagg ctaaccttct cattatcagt ggtgaacata gtgtcgtcac ctccaccttc	3540
gaacttcctt cctagatcgt aaagatagag gaaatcgtcc attgtaatct ccggggcaaa	3600
ggagatctct tttggggctg gatcactgct gggccttttg gttcctagcg tgagccagtg	3660
ggctttttgc tttggtgggc ttgttagggc cttagcaaag ctcttgggct tgagttgagc	3720
ttctcctttg gggatgaagt tcaacctgtc tgtttgctga cttgttgtgt acgctcagc	3780
tgctgctctt gcctctgtaa tagtggcaaa tttcttgtgt gcaactccgg gaacgcogtt	3840
tgttgccgcc tttgtacaac cccagtcac gtatataccg gcatgtggac cgttatacac	3900
aacgtagtag ttgatagag ggtgttgaat acccgattct gctctgagag gagcaactgt	3960
gctgttaagc tcagattttt gtgggattgg aattggatcg atctcgatcc cgcgaaatta	4020
atacactca ctataggag accacaacgg tttcctcta gaaataattt tgtttaactt	4080
taagaaggag atatacccat ggaaaagcct gaactcaccg cgacgtctgt cgagaagttt	4140
ctgatcgaag agttcgacag cgtctccgac ctgatgcagc tctcggaggg cgaagaatct	4200
cgtgctttca gcttcgatgt aggagggcgt ggatattgct tgccgggtaa tagctgcgcc	4260
gatggtttct acaaaagatc ttatgtttat cggcactttg catcggccgc gctcccatt	4320
ccggaagtgc ttgacattgg ggaattcagc gagagcctga cctattgcat ctcccgcgt	4380
gcacaggggtc tcacgttgca agacctgect gaaaccgaac tgcccogtgt tctgcagccg	4440
gtcgcggagg ctatggatgc gatcgtcgcg gccgatctta gccagacgag cggggtcggc	4500
ccattcggac cgcaaggaat cggtaatac actacatggc gtgatttcat atgocgatt	4560
gctgatcccc atgtgatca ctggcaaac gtgatggacg acaccgtcag tcgctccgtc	4620
gcgagggctc tcgatgagct gatgctttgg gccgaggact gcccgaagt ccggcacctc	4680
gtgcacgcgg atttcggctc caacaatgct ctgacggaca atggccgat aacagcggtc	4740
attgactgga gcgaggcgt gttcggggat tcccatacag aggtcgccaa catcttcttc	4800
tggaggccgt ggttggcttg tatggagcag cagacgcgct acttcgagcg gaggcacccg	4860
gagcttgcaag gatcgcgcg gctccggcg tatatgctcc gcattggtct tgaccaactc	4920
tatcagagct tggttgacgg caatttcgat gatgcagctt gggcgcaggg tcgatgcgac	4980
gcaatcgtcc gatecggagc cgggactgct gggcgtacac aaatgcgcc cagaagcgcg	5040
gccgtctgga ccgatggctg tgtagaagta ctcccgata gtggaaaccg acgcccagc	5100
actcgtccga gggcaagga atagttaggt acagcttggg tcgatccggc tgctaacaaa	5160
gcccgaagg aagctgagtt ggctgctgcc accgctgagc aataactagc ataaccctt	5220
ggggcctcta aacgggtctt gaggggtttt ttgctgaaag gaggaactat atccgatga	5280
tcgggcgcgc cgtcgacgga tccactagtt ctgagcggc ccgcgcgctc gacggatata	5340

-continued

---

atgagccgta aacaaagatg attaagtagt aattaatacg tactagtaaa agtggcaaaa	5400
gataacgaga aagaaccaat ttctttgcat tcggccttag cggaaggcat atataagctt	5460
tgattatfff atttagtgta atgatttcgt acaaccaaag catttattta gtactctcac	5520
acttgtgtcg cggccggccg ctacaggaac aggtggtggc ggcctcggc gcgctcgtac	5580
tgctccacga tgggtgtagc ctcggtgtgg gaggtgatgt ccagcttggg gtccacgtag	5640
tagtagccgg gcagctgcac gggcctcttg gccatgtaga tggacttggg ctccaccagg	5700
tagtggccgc cgtcctcag cttcagggcc ttgtggatct cgccctcag cacgcgctcg	5760
cgggggtaca ggcgctcggg ggaggcctcc cagcccatag tcttctctg cattacgggg	5820
ccgtcggagg ggaagtccac gccgatgaac ttcacctgt agatgaagga gccgctcgc	5880
agggaggagt cctgggtcac ggtcaccacg ccgccgtcct cgaagtccat cacgcctcc	5940
cacttgaagc cctcggggaa ggacagcttc ttgtagtcgg ggatgtcggc ggggtgcttc	6000
acgtacacct tggagccgta ctggaactgg ggggacagga tgtcccaggc gaaggcagc	6060
gggcccctc tggtcacctt cagcttggcg gtctgggtgc cctcgtaggg gcggccctcg	6120
cctcgcctc cgatctcgaa ctcggtggcg ttcacggagc cctccatgcg caccttgaag	6180
cgcatgaact ccttgatgac gtccctggag gaggccatgg gccgcttggg gggctatgga	6240
agactttctt agttagtgtg gtgaataagc aatgttggga gaatcgggac tacttatagg	6300
ataggaataa aacagaaaag tattaagtgc taatgaaata tttagactga taattaaat	6360
cttcacgtat gtccacttga tataaaaacg tcaggaataa aggaagtaca gtagaattta	6420
aagtgactct tttatataat acccgtgttc tctttttggc tagctagtgg cataaaaaat	6480
aatctatatt tttatcatta ttttaaatat cttatgagat ggtaaatatt taccataatt	6540
tttttacta ttatttatta ttgtgtgtg taatacatat agaagttaat tacaaaatft	6600
atttactttt tcattatttt gatatgattc accattaatt tagtgttatt atttataata	6660
gttcatttta atctttttgt atataattatg cgtgcagtac ttttttecta catataacta	6720
ctattacatt ttatttatat aatattttta ttaatgaatt ttcgtgataa tatgtaatat	6780
tgttcattat tatttcagat ttttataaaa tatttgtgtt attatttatg aaatagttaa	6840
tttttttagt atttgatttt atgatgataa agtgttctaa attcaaaaga agggggaaa	6900
cgtaaacatt aaaaaacgtc atcaaaacaa aacaaaatct tgtaataaaa gataaaactg	6960
tttgtttga tcactgttat ttcgtaatat aaaaacatta tttatattta tattgttgac	7020
aaccaaatft gcctatcaaa tctaaccaat ataatgcatg cgtggcaggt aatgtactac	7080
catgaactta agtcatgaca taataaacgg tgaatctgac caatgcatgt acctanctaa	7140
attgtatttg tgacacgaag caaatgattc aattcacaat ggagatggga acaaaataat	7200
gaagaacca gaactaagaa agcttttctg aaaaataaaa taaaggcaat gtcaaaagta	7260
tactgcatca tcagtcgaga aagcacatga tattttttta tcagtatcaa tgcagctagt	7320
tttattttac aatatcgata tagctagtft aatatattg cagctagatt tataaatatt	7380
tgtgttatta tttatcattt gtgtaatcct gtttttagta ttttagttaa tatatgatga	7440
taatgtattc caaatftaaa agaagggaaa taaattftaaa caagaaaaaa agtcatcaaa	7500
caaaaaacaa atgaaagggt ggaagatgt taccatgtaa tgtgaatgtt acagtatttc	7560
ttttattata gagttaacaa attaactaat atgattttgt taataatgat aaaaatfttt	7620
ttttattatt atttcataat ataaaaatag tttacttaat ataaaaaaa ttctatcgtt	7680
cacaacaaag ttggccacct aatttaacca tgcagtacc catggaccat attaggtaac	7740

-continued

---

catcaaacct gatgaagaga taaagagatg aagacttaag tcataacaca aaaccataaa	7800
aaacaaaaat acaatcaacc gtcaatctga ccaatgcatg aaaaagctgc aatagtgagt	7860
ggcgacacaa agcacatgat tttcttacia cggagataaa accaaaaaa tatttcatga	7920
acaacctaga acaataaag cttttatata ataaatatat aaataaataa aggctatgga	7980
ataatatact tcaatatatt tggattaaat aaattggttg cgggggtgat atatttatac	8040
acacctaaag tcacttcaat ctcatcttca cttaactttt attttttttt tctttttatt	8100
tatcataaag agaatttga taatatactt tttaacatat ttttatgaca ttttttattg	8160
gtgaaaactt attaaaaac ataaattttg taagttagat ttatttaaag agttcctctt	8220
cttattttta attttttaat aaatttttaa ataactaaaa tttgtgttaa aaatgttaa	8280
aaagtgtgtt attaacctt ctcttcgagg atccgtaccg agctcggatc cactagtaac	8340
ggccgccagt gtgctggaat tcaggtcctg caggtctact ctttacatgt tctttactcc	8400
gtctcaaaat ttcctttttt tgttggtctt ctccgaacga gttggagaaa tcgttaacct	8460
taatcgaaga tctagattcc tctacatacg tttgatctct ctctcagtat ggattacaaa	8520
gcgccaagga gatactactc acacggagtt gttgcgagac agcaagattt cgcaacagat	8580
atagttacga gaagaagacc ttatgtocct tacgaccgtc caaataagtt ttcaaggagt	8640
ctggtttgga cgtcaaaaga gtacaaatca cccgagggca ataatatgcc aaggaccaat	8700
gatgtgtcac cgaaacacc agtttttagt ttggcgagga agaattgctc ttgtgggcca	8760
atgagatctt ctagtctcag aaaatgggta tgtaagtatt ggaaagatgg aaagtgcaag	8820
aggggtgagc agtgcagtt cttacactct tggctctgtt tccctggatt ggccatggta	8880
gcttctcttg aagggcacia taaggaacta aaggggatcg ctctccctga gggttcagat	8940
aaactccttt cagtcagtat tgatggtaca ttgcgagttt gggactgcaa ttctggtcag	9000
tgtgtacatt ccatcaaac tgcgcagaa gcagggtctc taatcagtgaggcccttg	9060
gttttccttg gcttgccaaa cgctataaag gcttttaacg ttcaaaccag tcaagatttg	9120
catcttcaag cagcaggggt ggttggtcag gtgaatgcaa tgactattgc aaacggaatg	9180
ctttttgctg gaacaagttc tggtagtata ttagtctgga aagctactac agactctgag	9240
tctgatccat tcaaatactt gacatctctt gagggacata gtggtgaagt cacttgtttt	9300
gctgttgag gtcaaatgct atactctggt tctgtcgata aaacaatcaa gatgtgggat	9360
ctcaacaccc tgcaatgtat aatgacctg aagcaacata ccggcactgt cacttcactc	9420
ttatgttggg ataaatgttt gatatcgtct tccctggatg ggaccataaa agtttgggct	9480
tattctgaaa acggaatcct gaaagtgtt caaactcgca gacaagaaca gagtagtggt	9540
catgctcttt ctggtatgca tgatgcagaa gccaaaccga taatattctg ctcttaccaa	9600
aacggaacg ttggcatttt cgacctacca tcttttcaag aaagaggaag gatgttctct	9660
acgcacacga tcgccacact cacaattggt cctcaaggat tggattcag tggagacgag	9720
agtggttaact tgcgtgatg gaccttagct gctggcaaca aagtttagtc ttttcgacta	9780
aagaattctg atttaatttt gtggtttata tgttgagtta actgttaaga gagttttatt	9840
ttgtaatagg tgtatcagtc aataaacaat ctttgtatca accaaatgta atttttctcg	9900
ttaattcgat ttcagagttt ttactttaag ataaacaaac tctttcacac atcatttaat	9960
gaaagtggag aagcttaaaa aacaaacaaa gaaactgatc catttttggc gggctctctt	10020
ctactcttat tcatatgtgt taacgaacta tagcgtaaaa ttcagagcaa gcgatctccg	10080

-continued

---

```

atttgaacgt ggctatcacc ggaggccac cactacgggc gatacgcctc aagtgaggat 10140
taaagtgtct tgggtgtgac gttgaagaaa ctgcgccatg gtttttgta tctctgcagc 10200
caagtgtcgt tctttctctg ccacttctca tcaagctaca gtgaatttaa aaatggcgtc 10260
tttctttgat ctcgatatca taagctggat tggtttctta aacaaattcc tctccttttg 10320
ggtcttctgg gtttgccttg taagtgtttg tgtttttgcc tctgagaaaa aatcgcggcc 10380
gcaagatga actaaaatgc atgtaggtgt aagagctcat ggagagcatg gaatattgta 10440
tccgaccatg taacagtata ataactgagc tccatctcac ttcttctatg aataaacaaa 10500
ggatgttatg atatattaac actctatcta tgcaccttat tgttctatga taaatttcct 10560
cttattatta taaatcatct gaatcgtgac ggcttatgga atgcttcaaa tagtacaaaa 10620
acaaatgtgt actataagac tttctaaca attctaacct tagcattgtg aacgagacat 10680
aagtgttaag aagacataac aattataatg gaagaagttt gtctocattt atatattata 10740
tattaccacac ttatgtatta tattaggatg ttaaggagac ataacaatta taaagagaga 10800
agtttgtatc catttatata ttataacta cccatttata tattatactt atccacttat 10860
ttaatgtctt tataaggttt gatccatgat atttctaata ttttagtga tatgtatatg 10920
aaagggtact atttgaactc tcttactctg tataaaggtt ggatcatcct taaagtgggt 10980
ctatttaatt ttattgcttc ttacagataa aaaaaaatt atgagttggt ttgataaaat 11040
attgaaggat ttaaaataat aataaataac atataatata tgtatataaa tttattataa 11100
tataacattt atctataaaa aagtaaatat tgtcataaat ctatacaatc gtttagcctt 11160
gctggacgaa tctcaattat ttaaacgaga gtaaacatat ttgacttttt ggttatttaa 11220
caaattatta tttaacacta tatgaaattt tttttttat cagcaaagaa taaaattaaa 11280
ttaagaagga caatgggtgc ccaatcotta tacaaccaac tccacaaga aagtcaagtc 11340
agagacaaca aaaaaacaag caaaggaaat tttttaattt gagttgtctt gtttgctgca 11400
taatttatgc agtaaaacac tacacataac ccttttagca gtagagcaat ggttgaccgt 11460
gtgcttagct tcttttattt tattttttta tcagcaaaga ataaataaaa taaaatgaga 11520
cacttcaggg atgtttcaac aagctctaga gggcccaatt cgccctatag tgagtcgtat 11580
tacaattcac tggccgtcgt tttacaacgt cgtgactggg aaaaccctgg cgttacccaa 11640
cttaatcgcc ttgcagcaca tcccccttc gccagctggc gtaatagcga agaggccgcg 11700
accgatcgcc cttccaaca gttgcgcagc ctatac 11736

```

```

<210> SEQ ID NO 26
<211> LENGTH: 12929
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: vector pKR1167
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: (8496)..(8496)
<223> OTHER INFORMATION: n is a, c, g, or t

<400> SEQUENCE: 26

```

```

ggccgcaagt atgaactaaa atgcatgtag gtgtaagagc tcatggagag catggaatat 60
tgtatccgac catgtaacag tataataaact gagctccatc tcacttcttc tatgaataaa 120
caaaggatgt tatgatatat taacctctca tctatgcacc ttattgttct atgataaatt 180
tcctcttatt attataaatc atctgaatcg tgacggctta tggaaatgctt caaatagtac 240
aaaaacaaat gtgtactata agactttcta aacaattcta accttagcat tgtgaacgag 300

```

-continued

---

acataagtgt taagaagaca taacaattat aatggaagaa gtttgtctcc atttatatat	360
tatatattac ccacttatgt attatattag gatgttaagg agacataaca attataaaga	420
gagaagtttg tatccattta tatattatat actaccocatt tatatattat acttatccac	480
ttatttaatg tctttataag gtttgatcca tgatatttct aatatttttag ttgatatgta	540
tatgaaaggg tactatttga actctcttac tctgtataaa ggttgatca tccttaaagt	600
gggtctatth aatthttattg cttcttacag ataaaaaaaa aattatgagt tggthttgata	660
aaatattgaa ggattthaaa taataataaa taacatataa tatatgtata taaatthatt	720
ataatataac atthtatctat aaaaaagtaa atattgtcat aaatctatac aatcgthtag	780
ccttgctgga cgaatctcaa thattthaaac gagagthaac atatttgact thttgthtat	840
thaaacaaatt atthattthaac actatathgaa atththththth thtathcagcaa agaathaaaat	900
thaaathaga aggacathgg tgtcccaatc cttatataac caactthccac aagaaagthca	960
agthcagagac aacaaaaaaaa caagcaagg aaatthththth atthgagthg thctgththg	1020
thcathaatth atgathgthaa acactathaca thaacctthth agcathgagag caathgthg	1080
ccgthgthctt agctthctthth atththattth thtathcagca aagaathaaaat aaaaathaaaat	1140
gagacactthc aggathgthth caacaagthc thagagggccc aatthcgcctt athgthgagthc	1200
gthattacath thactggcgc thgthththaca acgthcgtgac thgggthaaac cthggcgtthc	1260
ccactthaat cgcctthgag cacathcccc ththcgcagc thggcghaata gchgaagagthc	1320
ccgchccgath cgcctthccc aacaththgcg chgcctathc thacgagathc cggcchggcca	1380
gathctgchag gathathcaag cthggcgcgc cgtthctathg thgthacctha atcgthathg	1440
thtgathacath aagththathg athaathgth gccgcgthct aacghacaata thgthcathg	1500
gthgactctc agthacaathc gthctgathc cgcathgthha agchcagcccc gathcchcgc	1560
aacacchcgc gathcgcchc gathgggctthg thctgthcccc gathcchcctt acathgacaag	1620
thgthaccgthc thccgggagct gathgthgth gagthththca cchgthathc chgaacgchc	1680
gathcghaaag ggcctchgth thcgcctath ththathgth aathgthath ccaaaathccc	1740
thaaacgthg ththctgthc actgagchc agacchcgha gaaaagathca aagghathctt	1800
thgathathct thththctgch gcghaathctg ctgctthgcaa acaaaaaac caccgthacc	1860
agchgthgth thgththcchg athcaagath accaathctt ththcghaag thactgthctt	1920
chcagagchc chgathacaa athactgthct thctathgth cchgaththg gchcchactt	1980
caagaathct gthacchcgc thacathacct cgcctgthha atcctgthac chgthgthc	2040
thcchagthgc gathagthct gthctthaccg gthgagthca agathgathg thcchgthaa	2100
ggcchcagchc thcgggthgaa chgggggthc gthcathcag cchcagctthg agchgaacgch	2160
thacchcgha cthgathathc thacagchgth gaththgath agchcchcgc thcchghaag	2220
gathaaagchc gathcagthath chgthaaagch chgggthcgha achgghagch gchcaggggha	2280
gctthcchggg gghaaacchc gthathctthha thgthctgthc gggththcgc acctctgact	2340
thgagchthg thththgthg gthcchcagch ggggchggch cthathgthaaa achcchcgha	2400
chcgggchctt thacgththc thggchctthg ctgchcctth gthcathathg thctthctgch	2460
gthathcccc gathctgthg athaacgthath thcchcctt gathgagchg athcchcchc	2520
chcchcchc chcagchcgc gchcchcagth agthgchcgh gathcchgha agchcchcaath	2580
achcchaaacch cctctcccc chcchgthgch gathaththha thcchggthg thcathctgch	2640

-continued

---

atcgatctag taacatagat gacaccgcgc gcgataatth atcctagttt gcgcgctata	2700
ttttgttttc tategcgat taaatgtata attgcgggac tctaatcata aaaaccatc	2760
tcataaataa cgtcatgcat tacatgtaa ttattacatg cttaacgtaa ttcaacagaa	2820
attatatgat aatcatcgca agaccggcaa caggattcaa tcttaagaaa ctttattgcc	2880
aatgtttga acgatctgct tcgacgcaact ccttctttag gtacctcact attcctttgc	2940
cctcggacga gtgctggggc gtcggtttcc actatcggcg agtacttcta cacagccatc	3000
ggtccagaac gccgcgcttc tgcggcgcat ttgtgtacgc ccgacagtcc cggtccgga	3060
tcggacgatt gcctcgcac gaccctgcgc ccaagctgca tcatcgaat tgccgtcaac	3120
caagctctga tagagttggt caagaccaat gcggagcata tacgcccga gccggcgga	3180
tctgcaagc tccggatgcc tccgctcga gtacgcgctc tgctgctcca tacaagcaa	3240
ccacggcctc cagaagaaga tgttggcgac ctcgtattgg gaatccccga acatcgctc	3300
gctccagtea atgaccgctg ttatgcggcc attgtcctgc aggacattgt tggagccgaa	3360
atccgcgtgc acgaggtgcc ggacttcggg gcagtcctcg gcccaaagca tcagctcatc	3420
gagagcctgc gcgacggacg cactgacggt gtcgtccatc acagtttgcc agtgatacac	3480
atggggatca gcaatcgcgc atatgaaac acgccatgta gtgtattgac cgattccttg	3540
cggtccgaat gggccgaacc cgctcgtctg gctaagatcg gccgcagcga tcgcatccat	3600
ggcctccgcg accggctgca gaacagcggg cagttcggtt tcaggcaggc cttgcaacgt	3660
gacaccctgt gcaaggcggg agatgcaata ggtcaggctc tcgctgaatt ccccaatgtc	3720
aagcacttcc ggaatcggga gcgcggccga tgcaaagtgc cgataaacat aacgatcttt	3780
gtagaaacca tcggcgcagc tatttaccgg caggacatat ccacgcctc ctacatcga	3840
gctgaaagca cgagattctt cgccctccga gagctgcac aggtcggaga cgctgtcga	3900
cttttcgatc agaaacttct cgacagacgt cgcggtgagt tcaggctttt tcatggttta	3960
ataagaagag aaaagagttc tttgttatg gctgaagtaa tagagaaatg agctcgagcg	4020
tgctctctcc aaatgaaatg aacttctta tatagaggaa gggctctgcg aaggatagtg	4080
ggattgtgcg tcattccctta cgtcagtgga gatgtccat caatccactt gctttgaaga	4140
cgtggttga acgtctctt tttccacgat gctcctcgtg ggtgggggtc catctttggg	4200
accactgtcg gcagaggcat cttgaatgat agcctttctt ttatcgcaat gatggcattt	4260
gtaggagcca ccttcctttt ctactgtcct ttcgatgaag tgacagatag ctgggcaatg	4320
gaatccgagg aggtttcccg aaattatcct ttgttgaaaa gtctcaatag ccctttggtc	4380
ttctgagact gtatctttga catttttga gtagaaccaga gtgtcgtgct ccaccatggt	4440
gacgaagatt ttcttctgt cattgagtcg taaaagactc tgatgaact gttcgcagc	4500
cttcacggcg agttctgtta gatcctcgat ttgaatctta gactccatgc atggccttag	4560
attcagtagg aactacctt tttagactc caatctctat tacttgctt ggtttatgaa	4620
gcaagccttg aatcgtccat actggaatag tacttctgat cttgagaaat atgtctttct	4680
ctgtgttctt gatgcaatta gtctcgaatc ttttgactgc atctttaaacc ttcttggaa	4740
ggtatttgat ctctggaga ttgttactcg ggtagatcgt cttgatgaga cctgctgct	4800
aggcctctct aaccatctgt gggtcagcat tctttctgaa attgaagagg ctaaccttct	4860
cattatcagt ggtgaacata gtgtcgtcac cttcacctc gaacttctt cctagatcgt	4920
aaagatagag gaaatcgtcc attgtaatct ccggggcaaa ggagatctct tttgggctg	4980
gatcactgct gggccttttg gttcctagcg tgagccagtg ggctttttgc tttggtggc	5040

-continued

---

ttgtagggc	cttagcaaag	ctcttgggct	tgagttgagc	ttctcctttg	gggatgaagt	5100
tcaacctgtc	tgtttgctga	cttggtgtgt	acgcgtcagc	tgetgctctt	gcctctgtaa	5160
tagtggaaca	tttcttgtgt	gcaactcccg	gaacgcctgt	tggtgccgcc	ttgtacaac	5220
cccagtcac	gtatataccg	gcatgtggac	cggtatacac	aacgtagtag	ttgatatgag	5280
ggtgttgaat	acccgattct	gctctgagag	gagcaactgt	gctgttaagc	tcagattttt	5340
gtgggattgg	aattggatcg	atctcgatcc	cgcgaaatta	atacgactca	ctatagggag	5400
accacaacgg	tttcctcta	gaaataattt	tgtttaactt	taagaaggag	atatacccat	5460
ggaaaagcct	gaactcaacc	cgacgtctgt	cgagaagttt	ctgatcgaaa	agttcgacag	5520
cgtctccgac	ctgatgcagc	tctcggaggg	cgaagaatct	cgtgctttca	gcttcgatgt	5580
aggagggcgt	ggatatgtcc	tgccggtaaa	tagctgcgcc	gatggtttct	acaaagatcg	5640
ttatgtttat	cggcactttg	catcggccgc	gctcccgatt	ccggaagtgc	ttgacattgg	5700
ggaattcagc	gagagcctga	cctattgcat	ctcccgcctg	gcacagggtg	tcacgttgca	5760
agacctgcct	gaaaccgaac	tgcccgtctg	tctgcagccg	gtcgcggagg	ctatggatgc	5820
gatcgtcgcg	gcccgatctta	gccagacgag	cggttcggc	ccattcggac	cgcaaggaaat	5880
cggtaaatca	actacatggc	gtgatttcat	atgcgcgatt	gctgatcccc	atgtgtatca	5940
ctggcaaaact	gtgatggacg	acaccgtcag	tgccgtccgtc	gcgcaggctc	tcgatgagct	6000
gatgctttgg	gcccaggact	gccccgaagt	ccggcacctc	gtgcacgcgg	atttcggctc	6060
caacaatgtc	ctgacggaca	atggccgcat	aacagcggtc	attgactgga	gagaggcgat	6120
gttcggggat	tcccataacg	aggctgccaa	catcttcttc	tggaggccgt	ggttggttg	6180
tatggagcag	cagacgcgct	acttcgagcg	gaggcatccg	gagcttgca	gatcgcgcg	6240
gctccgggcg	tatatgtccc	gcattggctc	tgaccaactc	tatcagagct	tggttgacgg	6300
caatttcgat	gatgcagctt	gggcgcaggg	tcgatgcgac	gcaatcgtcc	gatccggagc	6360
cgggactgtc	ggcgctacac	aaatcgcccg	cagaagcgcg	gccgtctgga	ccgatggctg	6420
tgtagaagta	ctcgcgata	gtggaaaaccg	acgccccagc	actcgtccga	gggcaaaagga	6480
atagtgaggt	acagcttggg	tcgatccggc	tgctaacaaa	gcccgaagg	aagctgagtt	6540
ggctgctgcc	accgctgagc	aataactagc	ataaccctt	gggcctcta	aacgggtctt	6600
gaggggtttt	ttgctgaaag	gaggaactat	atccggatga	tcgggcgcgc	cgctcgacgga	6660
tccactagtt	ctagagcggc	ccgcgcctgc	gacggatata	atgagccgta	aacaaagatg	6720
attaagtagt	aattaatacg	tactagtaaa	agtggcaaaa	gataacgaga	aagaaccaat	6780
ttctttgcat	tcggccttag	cggaaaggcat	atataagctt	tgattatttt	atttagtgta	6840
atgatttcgt	acaaccaaag	catttattta	gtactctcac	acttggtctg	cggccggccg	6900
ctacaggaac	aggtggtggc	ggccctcggc	gcgctcgtac	tgctccacga	tggtgtagtc	6960
ctcgttgtgg	gaggtgatgt	ccagcttggg	gtccacgtag	tagtagccgg	gcagctgcac	7020
gggcttcttg	gccatgtaga	tggacttgaa	ctccaccagg	tagtggccgc	cgctcctcag	7080
cttcagggcc	ttgtggatct	cgccctcag	cacgcctcgc	cggggtaca	ggcgtcgggt	7140
ggaggcctcc	cagcccatag	tcttctctg	cattacgggg	ccgtcggagg	ggaagttcac	7200
gccgatgaac	ttcaccttgt	agatgaagga	gccgtcctgc	agggaggagt	cctgggtcac	7260
ggtcaccacg	ccgcctcct	cgaagtcat	cacgcctcgc	cacttgaagc	cctcggggaa	7320
ggacagcttc	ttgtagtcgg	ggatgtcggc	gggtgcttc	acgtacacct	tgagccgta	7380

-continued

---

ctggaactgg ggggacagga tgtcccagge gaagggcagg gggccgccct tggtcacctt	7440
cagcttggcg gtctgggtgc cctcgtaggg gcggccctcg ccctcgcctt cgatctcgaa	7500
ctcgtggcgg ttcacggagc cctccatgcg caccttgaag cgcctgaact ccttgatgac	7560
gtcctcggag gaggccatgg gccgcttggg gggctatgga agactttctt agttagtgtt	7620
gtgaataagc aatgttggga gaatcgggac tacttatagg ataggaataa aacagaaaag	7680
tattaagtgc taatgaaata tttagactga taattaaaat cttcacgtat gtccacttga	7740
tataaaaacg tcaggaataa aggaagtaca gtagaattta aaggtactct tttatataat	7800
accctgttcc tctttttggc tagctagtgg cataaaaaat aatctatatt tttatcatta	7860
ttttaaatat cttatgagat ggtaaatatt tatcataatt ttttttacta ttatttatta	7920
tttgtgtgtg taatacatat agaagttaat tacaaaatctt atttactttt tcattatctt	7980
gatatgattc accattaatt tagtgttatt atttataata gttcatttta atctttttgt	8040
atataattatg cgtgcagtac ttttttctca catataacta ctattacatt ttatttataat	8100
aatattttta ttaatgaatt ttcgtgataa tatgtaatat tggtcattat tatttcagat	8160
tttttaaaaa tatttgtgtt attatttatg aaatatgtaa tttttttagt atttgatttt	8220
atgatgataa agtgttctaa attcaaaaga aggggggaaag cgtaaacatt aaaaaacgtc	8280
atcaaacaaa aacaaaatct tgtaataaaa gataaaactg tttgttttga tcaactgttat	8340
ttcgtaatat aaaaacatta tttatattta tattgttgac aaccaaattt gcctatcaaa	8400
tctaaccaat ataatgcatg cgtggcaggt aatgtactac catgaactta agtcatgaca	8460
taataaacgg tgaatctgac caatgcatgt acctanctaa attgtatttg tgacacgaag	8520
caaatgattc aattcacaat ggagatggga aacaaaatat gaagaacca gaactaagaa	8580
agcttttctg aaaaaataaa taaaggcaat gtcaaaagta tactgcatca tcagtccaga	8640
aagcacatga tattttttta tcagtatcaa tgcagctagt tttattttac aatatcgata	8700
tagctagtctt aaatatattg cagctagatt tataaatatt tgggttatta tttatcattt	8760
gtgtaactct gtttttagta ttttagttta tatatgatga taatgtattc caaatttaaa	8820
agaagggaaa taaatttaaa caagaaaaaa agtcatcaaa caaaaaacaa atgaaagggg	8880
ggaaagatgt taccatgtaa tgtgaatgtt acagtatttc tttatttata gagttaacaa	8940
attaactaat atgattttgt taataatgat aaaatatttt tttatttatt atttcataat	9000
ataaaaatag tttacttaat ataaaaaaaa ttctatcgtt cacaacaaag ttggccacct	9060
aatttaacca tgcattgacc catggaccat attaggtaac catcaaacct gatgaagaga	9120
taaagagatg aagacttaag tcataacaca aaaccataaa aaacaaaaat acaatcaacc	9180
gtcaatctga ccaatgcatg aaaaagctgc aatagtgagt ggcgacacaa agcacatgat	9240
tttcttacia cggagataaa accaaaaaaa tatttcatga acaacctaga acaataaag	9300
cttttatata ataaatataat aaataaataa aggctatgga ataataact tcaatatatt	9360
tggattaaat aaattgttgg cgggggtgat atatttatac acacctaaag tcacttcaat	9420
ctcattttca ctttaactttt attttttttt tctttttatt tatcataaag agaatttga	9480
taataatactt tttaacatat ttttatgaca ttttttattg gtgaaaactt attaaaaatc	9540
ataaattttg taagttagat ttattttaaag agttcctctt cttattttta attttttaat	9600
aaatttttaa ataactaaaa tttgtgttaa aaatgttaa aaagtgtgtt attaacctt	9660
ctcttcgagg atccgtaccg agctcggatc cactagtaac ggcgccaggt gtgctggaat	9720
tcaggctctg caggtctact ctttacatgt tctttactcc gtctcaaaaat ttcctttttt	9780

-continued

---

tgttggctct	ctccgaacga	gttggagaaa	tcgttaaccc	taatogaaga	tctagattcc	9840
tctacatacg	tttgatctct	ctctcagtat	ggattacaaa	gcgccaagga	gatactactc	9900
acacggagtt	gttgcgagac	agcaagattt	cgcaacagat	atagttacga	gaagaagacc	9960
ttatgtccct	tacgaccgtc	caaataagtt	ttcaaggagt	ctggtttggg	cgtaaaaaga	10020
gtacaaatca	cccgagggca	ataatatgcc	aaggaccaat	gatgtgtcac	cgaaccacc	10080
agtttttagt	ttggcgagga	agaatgctgc	ttgtgggcca	atgagatctt	ctagtctcag	10140
aaaatgggta	tgtaagtatt	ggaaagatgg	aaagtgcaag	aggggtgagc	agtgccagtt	10200
cttacctct	tggctctgtt	tccctggatt	ggccatggta	gcttctcttg	aagggcacia	10260
taaggaacta	aaggggatcg	ctctccctga	gggttcagat	aaactctttt	cagtcagtat	10320
tgatggtaca	ttgcgagttt	gggactgcaa	ttctggtcag	tgtgtacatt	ccatcaacct	10380
tgacgcagaa	gcagggctct	taatcagtga	agggccttgg	gttttcttgg	gcttgccaaa	10440
cgctataaag	gcttttaacg	ttcaaaccag	tcaagatttg	catcttcaag	cagcaggggt	10500
ggttggctag	gtgaatgcaa	tgactattgc	aaacggaatg	ctttttgctg	gaacaagttc	10560
tggtagtatc	ttagtctgga	aagctactac	agactctgag	tctgatccat	tcaaatactt	10620
gacatctctt	gagggacata	gtgggtgaag	cacttgtttt	gctgttggag	gtcaaatgct	10680
atactctggg	tctgtcgata	aaacaatcaa	gatgtgggat	ctcaacaccc	tgcaatgtat	10740
aatgaccctg	aagcaacata	ccggcactgt	cacttcactc	ttatgttggg	ataaatgttt	10800
gatatcgtct	tccttgatg	ggaccataaa	agtttgggct	tattctgaaa	acggaatctt	10860
gaaagtgtgt	caaactcgca	gacaagaaca	gagtagtggt	catgctcttt	ctggtagtga	10920
tgatgcagaa	gccaaccoga	taatattctg	ctcttaccaa	aacggaaccg	ttggcatttt	10980
cgacctacca	tcttttcaag	aaagaggaag	gatgttctct	acgcacacga	tcgccacact	11040
cacaattggg	cctcaaggat	tgttattcag	tggagacgag	agtggttaact	tgcgtgtatg	11100
gaccttagct	gctggcaaca	aagtttagtc	ttttcgacta	aagaattctg	atttaatttt	11160
gtggtttata	tgttgagtta	actgttaaga	gagttttatt	ttgtaatagg	tgtatcagtc	11220
aataaacaat	ctttgtatca	accaaaatgta	atttttctcg	ttaattcgat	ttcagagttt	11280
ttactttaag	ataaacaac	tctttcacac	atcatttaat	gaaagtggag	aagcttaaaa	11340
aacaaacaaa	gaaactgatc	catttttggc	gggtcttctt	ctactcttat	tcatatgtgt	11400
taacgaacta	tagcgtaaaa	ttcagagcaa	gcgatctccg	atgtgaacgt	ggctatcacc	11460
ggaggcccac	cactacgggc	gatacgtct	aagtgggat	taaagtgctc	tggtggtgac	11520
gttgaagaaa	ctcggccatg	gtttttggtt	tctctgcagc	caagtgtcgt	tctttctctg	11580
ccacttctca	tcaagctaca	gtgaatttaa	aaatggcgtc	tttctttgat	ctcgatatac	11640
taagctggat	tggtttctta	aacaaattcc	tctccttttg	ggtcttctgg	gtttgccctg	11700
taagtgtttg	tgtttttgcc	tctgagaaaa	aatcggcgcc	gcatggagag	atctcaacgg	11760
cagtctctct	cgccaccgtc	gccgtctctc	tectctctct	ccgtctctcg	ggacaccgtc	11820
ctcgtccctc	ccgaaaagag	gcggagggcg	gcgacggcca	aggccggcgc	cgagcctaat	11880
aagaggatcc	gcaaggaccc	cgccggcgcc	gcccggggga	agaggagctc	cgtctacag	11940
ggagtcacca	ggcacagggt	gacgggcagg	ttcgaggcgc	atctctggga	caagcactgc	12000
ctcggcgcgc	tccacaacaa	gaagaaaggc	aggcaagtct	acctgggggc	gtatgacagc	12060
gaggaggcag	ctgctctgtc	ctatgacctc	gcagctctca	agtactgggg	tcctgagact	12120

-continued

---

```

ctgctcaact tccctgtgga ggattactcc agcgagatgc cggagatgga ggcctgtgcc 12180
cgggaggagt acctggctc cctccgccgc aggagcagcg gcttctccag gggcgtctcc 12240
aagtacagag gcgtgccag gcatcaccac aacgggaggt gggaggcacg gattgggcga 12300
gtctttggga acaagtaoct ctacttggga acatttgaca ctcaagaaga ggcagccaag 12360
gcctatgacc ttgcggccat tgaataccgt ggcgtcaatg ctgtaaccaa cttcgacatc 12420
agctgctacc tggaccaccc gctgttctcg gcacagctcc aacaggagcc acaggtggtg 12480
ccggcactca accaagaacc tcaacctgat cagagcgaaa ccggaactac agagcaagag 12540
ccggagtcaa gcgaagccaa gacaccggat ggcagtgcag aaccgatga gaacgcggtg 12600
cctgacgaca ccgcgagacc cctcaccaca gtcgacgaca gcatcgaaga gggcttgtgg 12660
agcccttgca tggattacga gctagacacc atgtcgagac caaactttgg cagctcaatc 12720
aatctgagcg agtggttcgc tgacgcagac ttcgactgca acatcggatg cctgttcgat 12780
gggtgttctg ccgctgacga aggaagcaag gatggtgtag gtctggcaga tttcagctg 12840
tttgaggcag gtgatgtcca gctgaaggat gttctttcgg atatggaaga ggggatataa 12900
cctccagcga tgatcagtgt gtgcaacgc 12929

```

&lt;210&gt; SEQ ID NO 27

&lt;211&gt; LENGTH: 13268

&lt;212&gt; TYPE: DNA

&lt;213&gt; ORGANISM: Artificial Sequence

&lt;220&gt; FEATURE:

&lt;223&gt; OTHER INFORMATION: vector pKR92

&lt;400&gt; SEQUENCE: 27

```

cgcgctcga gtggcggat cccccggct gcaggaattc actggccgtc gttttacaac 60
gtcgtgactg ggaaaacct ggcgttacce aacttaatcg ccttgacga catccccctt 120
tcgccagctg gcgtaatagc gaagaggccc gcaccgatcg cccttccaa cagttgcgca 180
gcctgaatgg cgaatggatc gatccatcgc gatgtacett ttgttagtca gcctctcgat 240
tgctcatcgt cattacacag taccgaagtt tgatcgatct agtaacatag atgacaccgc 300
gcgcgataat ttatcctagt ttgcgcgcta tattttgttt tctatcgcgt attaatgta 360
taattgcggg actctaata taaaaacca tctcataaat aacgtcatgc attacatgtt 420
aattattaca tgcttaacgt aattcaacag aaattatatg ataatcatcg caagaccggc 480
aacaggatc aatcttaaga aactttattg ccaaatgttt gaacgatctg cttcgacgca 540
ctccttcttt actccacat ctcgtcctta ttgaaaacgt gggtagcacc aaaacgaatc 600
aagtcgctgg aactgaagtt accaatcacg ctggatgatt tgccagttgg attaactctg 660
cctttcccg catgaataat attgatgaat gcatgcgtga gggtagttc gatggtggca 720
atagctgcaa ttgccgcgac atcctccaac gagcataatt cttcagaaaa atagcagatg 780
tccatgttgt cagggcatgc atgatgcacg ttatgaggtg acggtgctag gcagatttcc 840
ctcaaagttt catagtcagt atcatattca tcattgcatt cctgcaagag agaattgaga 900
cgcaatccac acgctgcggc aaccttccgg cgttcgtggt ctatttgctc ttggacgttg 960
caaacgtaag tgttggatcg atccggggtg ggcgaagaac tccagcatga gatccccgcg 1020
ctggaggatc atccagccgg cgtccccgaa aacgattccg aagcccaacc tttcatagaa 1080
ggcggcgggt gaatcgaat ctcgtgatgg caggttgggc gtcgcttggg cggtcatttc 1140
gaaccccaga gtcccgtca gaagaactcg tcaagaaggc gatagaaggc gatgcgctgc 1200
gaatcgggag cggcgatacc gtaaagcacg aggaagcggg cagcccattc gccgccaagc 1260

```

-continued

---

tcttcagcaa	tateacgggt	agccaacgct	atgtcctgat	ageggtcgc	cacacccagc	1320
cggccacagt	cgatgaatcc	agaaaagcgg	ccattttcca	ccatgatatt	cggcaagcag	1380
geatcgccat	gggtcaecgac	gagatcctcg	ccgtcgggca	tgcgcgcctt	gagcctggcg	1440
aacagttecg	ctggcgcgag	cccctgatgc	tettegtcca	gatcatcctg	atcgacaaga	1500
ccggcttcca	tccgagtacg	tgctcgcctg	atgctgatgt	tcgcttggtg	gtcgaatggg	1560
caggtagccg	gatcaagcgt	atgcagccgc	cgcattgcat	cagccatgat	ggatactttc	1620
tcggcaggag	caaggtgaga	tgacaggaga	tectgccccg	gcacttcgcc	caatagcagc	1680
cagtcccttc	ccgcttcagt	gacaacgtcg	agcacagctg	cgcaaggaac	gcccgctcgtg	1740
gccagccacg	atagccgcgc	tgccctcgtc	tgcaagttcat	tcagggcacc	ggacaggtcg	1800
gtcttgacaa	aaagaaccgg	gccccctcgc	gctgacagcc	ggaacacggc	ggcatcagag	1860
cagccgattg	tctgtttgtc	ccagtcatag	ccgaatagcc	tctccacca	agcggccgga	1920
gaacctgcgt	gcaatccatc	ttgttcaatc	atgcaaacg	atccccgcaa	gcttgagac	1980
tggtgatttc	agcgtgtcct	ctccaaatga	aatgaacttc	cttatataga	ggaagggctc	2040
tgcaaggat	agtgggattg	tgctcctcc	cttacgtcag	tggagatata	acatcaatcc	2100
acttgctttg	aagacgtgg	tggaacgtct	tctttttcca	cgatgctcct	cgtgggtggg	2160
ggtccatctt	tgggaccact	gtcggcagag	geatcttcaa	cgatggcctt	tcctttatcg	2220
caatgatggc	attttagga	gccaccttc	ttttccacta	tcttcacaat	aaagtgacag	2280
atagctgggc	aatggaatcc	gaggagggtt	ccggatatta	ccctttgttg	aaaagtctca	2340
attgcccttt	ggtcttctga	gactgtatct	ttgatatttt	tggagtagac	aagcgtgctg	2400
tgctccacca	tggtgacgaa	gattttcttc	ttgtcattga	gtcgtaaagag	actctgtatg	2460
aactgttcgc	cagtctttac	ggcagattct	gttaggtcct	ctatttgaat	ctttgactcc	2520
atggcctttg	attcagtggt	aactaccttt	ttagagactc	caatctctat	tacttgctct	2580
ggtttgtaa	gcaagccttg	aatcgtccat	actggaatag	tacttctgat	cttgagaaat	2640
atatctttct	ctgtgtctt	gatgcagtta	gtcctgaatc	ttttgactgc	atctttaacc	2700
ttcttgggaa	ggatattgat	ctcctggaga	ttattgctcg	ggtagatcgt	cttgatgaga	2760
cctgctgcgt	aagcctctct	aaccatctgt	gggttagcat	tctttctgaa	attgaaaagg	2820
ctaactctct	cattatcagt	ggtgaacatg	gtatcgtcac	cttctccgtc	gaacttcctg	2880
actagatcgt	agagatagag	gaagtcgtcc	attgtgatct	ctggggcaaa	ggagtctgaa	2940
ttaattcgat	atggtggatt	tatcacaat	gggacccgcc	gccgacagag	gtgtgatgtt	3000
aggccaggac	tttgaatatt	tgccgcaacta	tcgtatagtg	gccgacaaat	tgacgccgag	3060
ttgacagact	gcctagcatt	tgagtgaatt	atgtgaggta	atgggctaca	ctgaattggt	3120
agctcaaaact	gtcagtattt	atgtatatga	gtgtatattt	tcgcataatc	tcagaccaat	3180
ctgaagatga	aatgggtatc	tggaatggc	gaaatcaagg	catcgatcgt	gaagtttctc	3240
atctaagccc	ccatttggac	gtgaatgtag	acacgtcgaa	ataaagattt	ccgaattaga	3300
ataatttggt	tattgctttc	gcctataaat	acgacggatc	gtaatttgct	gttttatcaa	3360
aatgtacttt	cattttataa	taacgctcgc	gacatctaca	tttttgaatt	gaaaaaaaaat	3420
tggaatttac	tctttctttt	tctccatatt	gaccatcata	ctcattgctg	atccatgtag	3480
atttcccgga	catgaagcca	tttacaattg	aatatatcct	gccgcocctg	ccgctttgca	3540
cccgtggag	cttgcatggt	ggtttctacg	cagaactgag	ccggttaggc	agataatctc	3600

-continued

---

cattgagaac	tgagccatgt	gcaccttccc	cccaacacgg	tgagcgacgg	ggcaacggag	3660
tgatccacat	gggactttta	aacatcatcc	gtcggatggc	gttgcgagag	aagcagtcga	3720
tccgtgagat	cagccgacgc	accgggcagg	cgcgcaacac	gatcgaaaag	tatttgaacg	3780
caggtacaat	cgagccgacg	ttcacgcgga	acgaccaagc	aagctagctt	taatgcggta	3840
gtttatcaca	gttaaattgc	taacgcagtc	aggcaccgtg	tatgaaatct	aacaatgcgc	3900
tcategteat	cctcggcacc	gtcaccctgg	atgctgtagg	cataggcttg	gttatgccgg	3960
tactgccggg	cctcttgogg	gatatcgtcc	attccgacag	catcgccagt	cactatggcg	4020
tgctgctagc	gctatatgcg	ttgatgcaat	ttctatgcgc	accggttctc	ggagcactgt	4080
ccgaccgctt	tggccgcgcg	ccagtcctgc	tcgcttcgct	acttgagacc	actatcgact	4140
acgcgatcat	ggcgaccaca	cccgtcctgt	ggtccaaccc	ctccgctgct	atagtgcagt	4200
cggtcttctga	cgttcagtcg	agccgtcttc	tgaaaacgac	atgtgcgaca	agtcctaagt	4260
tacgcgacag	gctgccgccc	tgccttttcc	ctggcgtttt	cttgctcgct	gttttagtcg	4320
cataaagtag	aatacttgcg	actagaaccg	gagacattac	gccatgaaca	agagcggcgc	4380
cgctggcctg	ctgggctatg	cccgcgtcag	caccgacgac	caggacttga	ccaaccaacg	4440
ggccgaactg	cacgcggcgc	gctgcaccaa	gctgttttcc	gagaagatca	ccggcaccag	4500
gcgcgaccgc	ccggagctgg	ccaggatgct	tgaccaccta	cgccctggcg	acgttgtgac	4560
agtgaccagg	ctagaccgcc	tggcccgcag	cacccgcgac	ctactggaca	ttgccgagcg	4620
catccaggag	gccggcgccg	gctcgcgtag	cctggcagag	ccgtgggccc	acaccaccac	4680
gccggcccgc	cgcatggtgt	tgaccgtgtt	cgccggcatt	gccgagttcg	agcgttcctc	4740
aatcatcgac	cgcacccgga	gcggggcgca	ggccgccaaag	gcccgaggcg	tgaagttagg	4800
ccccgcacct	acctcaacc	cggcacagat	cgcgcaacgc	cgcgagctga	tcgaccagga	4860
aggccgcacc	gtgaaagagg	cggtgcact	gcttggcgtg	catcgctcga	cctgtaccg	4920
cgacttgag	cgcagcgagg	aagtgcgcc	caccgaggcc	aggcggcgcg	gtgccttccg	4980
tgaggacgca	ttgaccgagg	ccgacgccct	ggcggccgcc	gagaatgaac	gccaagagga	5040
acaagcatga	aaccgcacca	ggacggccag	gacgaaccgt	ttttcattac	cgaagagatc	5100
gaggcggaga	tgatcgcggc	cggttacgtg	ttcgagccgc	ccgcgcacgt	ctcaaccgtg	5160
cggtgcgatg	aaatcctggc	cggtttgtct	gatgccaaag	tggcggcctg	gccggccagc	5220
ttggccgctg	aagaaaccga	gcgccgccgt	ctaaaaaggt	gatgtgtatt	tgagtaaac	5280
agcttgctgc	atgcggctcg	tgcgtatatg	atgcgatgag	taaataaaca	aatacgaag	5340
ggaacgcatg	aagttatcgc	tgtacttaac	cagaaaggcg	ggtcaggcaa	gacgaccatc	5400
gcaaccatc	tagcccgcgc	cctgcaactc	gccggggccg	atgttctggt	agtcgattcc	5460
gatccccagg	gcagtgcccg	cgattgggcg	gccgtgcccg	aagatcaacc	gctaaccgtt	5520
gtcggcatcg	accgcccgac	gattgaaccg	gacgtgaagg	ccatcggccg	gcgcgacttc	5580
gtagtgatcg	acggagcgcc	ccaggcggcg	gacttggctg	tgtccgcgat	caaggcagcc	5640
gacttcgtgc	tgattccggg	gcagccaagc	ccttacgaca	tatgggccac	cgccgacctg	5700
gtggagctgg	ttaagcagcg	cattgaggtc	acggatggaa	ggctacaagc	ggcctttgtc	5760
gtgtcggcgg	cgatcaaaag	cacgcgcata	ggcggtgagg	ttgccgaggc	gctggccggg	5820
tacgagctgc	ccattcttga	gtcccgatc	acgcagcgcg	tgagctaccc	aggcaactgc	5880
gccgcccgca	caaccgttct	tgaatcagaa	cccgagggcg	acgctgcccg	cgaggtccag	5940
gcgctggccg	ctgaaattaa	atcaaaactc	atgtgagtta	atgaggtaaa	gagaaaaatga	6000

-continued

---

gcaaaagcac aaacacgcta agtgccggcc gtccgagcgc acgcagcagc aaggctgcaa 6060  
cgttggccag cctggcagac acgccagcca tgaagcgggt caactttcag ttgccggcgg 6120  
aggatcacac caagctgaag atgtacgcgg tacgccaagg caagaccatt accgagctgc 6180  
tatctgaata catcgcgag ctaccagagt aatgagcaa atgaataaat gagtagatga 6240  
atthtagcgg ctaaaggagg cggcatggaa aatcaagaac aaccaggcac cgacgccgtg 6300  
gaatgcccc a tgtgtggagg aacgggagggt tggccaggcg taagcggctg ggttgtctgc 6360  
cggccctgca atggcactgg aacccccaaag cccgaggaat cggcgtgagc ggtcgcaaac 6420  
catccggccc ggtacaaaac ggcgcggcgc tgggtgatga cctggtggag aagttgaagg 6480  
ccgcgcaggg cgcacagcgg caacgcacgc aggcagaagc acgccccgtt gaatcgtggc 6540  
aagcggccgc tgatcgaac cgcacaagaat cccgcaacc gccggcagcc ggtgcgccgt 6600  
cgattagaa gccgcccaag ggcgacgagc aaccagattt ttctgttccg atgctctatg 6660  
acgtggggcac ccgcgatagt cgcagcatca tggacgtggc cgthttccgt ctgtcgaagc 6720  
gtgaccgacg agctggcgag gtgatccgct acgagcttcc agacgggcaac gtagaggttt 6780  
ccgcagggcc gccccgcatg gccagtggtg gggattacga cctggtactg atggcggttt 6840  
cccatcctaac cgaatccatg aaccgatacc ggaagggaa gggagacaag cccggccgcg 6900  
tgthccgtcc acacgttgcg gacgtactca agthctgcgg cgcagccgat ggcggaaagc 6960  
agaaagacga cctggtagaa acctgcattc ggttaaacac cacgcacgth gccatgcagc 7020  
gtacgaagaa ggccaagaac ggcgcctcgg tgacggatc cgaggggtgaa gccttgatta 7080  
gccgctacaa gatcgtaag agcgaaaccc ggcggccgga gtacatcgag atcgagctag 7140  
ctgattggat gtaccgcgag atcacagaag gcaagaaccc ggacgtgctg acggttcacc 7200  
ccgattactt thtgatgat cccggcatcg gccgthttct ctaccgctg gcacgccgcg 7260  
ccgcagggcaa ggcagaagcc agatggtgtg tcaagacgat ctacgaacgc agtggcagcg 7320  
ccggagagth caagaagthc tgtthcaccg tgcgcaagct gatcgggtca aatgacctgc 7380  
cggagtacga ththgaaggag gagggggggc aggctggccc gatcctagth atgcgctacc 7440  
gcaacctgat cgaaggcgaa gcatccgccg gthcctaatg tacggagcag atgctagggc 7500  
aaattgccct agcaggggaa aaaggtcgaa aaggtctctt tcctgtggat agcacgtaca 7560  
thgggaaccc aaagccgtac atthggaaac ggaacccgta cattgggaaac ccaaagccgt 7620  
acattgggaa ccggtcacac atgtaagtga ctgatataaa agagaaaaaa ggcgatttht 7680  
ccgcctaaaa ctctthaaaa ctthataaaa ctctthaaac ccgcctggcc thtgcataac 7740  
thgtctggcca gcgcacagcc gaagagctgc aaaaagccc taccttcgg thcgtcgcgt 7800  
ccctacgcc ccgcgcttcg cgtcggccta tccggccgc tggccgctca aaaaagggctg 7860  
gcctacggcc aggcaatcta ccagggcgcg gacaagccc gccgtcgcca ctgcacgcc 7920  
ggcgcaccaca tcaaggcacc ctgcctcgcg cgtthcggth atgacggthg aaacctctga 7980  
cacatgcagc thccggagac ggtcacagct thgtctgaag cggatgccgg gagcagacaa 8040  
gcccgtcagg gcgcgtcagc gggthgtggc gggthgtcgg gcgcagccat gacctagthc 8100  
cgtagcgata gcggagthg tactggctta actatgcggc atcagagcag atthgactga 8160  
gagthcacca thtgcgthgt gaaataccgc acagatgcgt aaggagaaaa tacgcacatc 8220  
ggcgtctctc cgtthcctcg ctactgact cgtcgcgctc ggtcgttcgg ctgcggcgag 8280  
cggthacagc thactcaag gcgthaaac ggtthaccac agaathcagg gataacgcag 8340

-continued

---

gaaagaacat	gtgagcaaaa	ggccagcaaa	aggccaggaa	ccgtaaaaag	gccgcgttgc	8400
tggcgTTTTT	ccataggctc	cgccccctg	acgagcatca	caaaaatcga	cgctcaagtc	8460
agagtgggcg	aaacccgaca	ggactataaa	gataaccaggc	gtttccccct	ggaagctccc	8520
tcgtgcgctc	tcctgttccg	accctgccc	ttaccggata	cctgtccgcc	ttctccctt	8580
cgggaagcgt	ggcgctttct	catagctcac	gctgtaggta	tctcagttcg	gtgtaggtcg	8640
ttcgctccaa	gctgggctgt	gtgcacgaac	cccccgttca	gcccgaccgc	tgcgccttat	8700
ccggtaaacta	tcgtcttgag	tccaaccocg	taagacacga	cttatcgcca	ctggcagcag	8760
ccactggtaa	caggattagc	agagcgaggt	atgtaggcgg	tgctacagag	ttcttgaagt	8820
ggtggcctaa	ctacggctac	actagaagga	cagtatttgg	tatctgcgct	ctgctgaagc	8880
cagttacctt	cggaaaaaga	gttggtagct	cttgatccgg	caaaaaacc	accgctggta	8940
gcggtggttt	ttttgtttgc	aagcagcaga	ttacgcgcag	aaaaaaagga	tctcaagaag	9000
atcctttgat	cttttctacg	gggtctgacg	ctcagtgtaa	cgaaaaactca	cgtaagggga	9060
ttttggctcat	gagattatca	aaaaggatct	tcacctagat	ccttttaaat	taaaaatgaa	9120
gttttaaatc	aatctaaagt	atatatgagt	aaacttggtc	tgacagttac	caatgcttaa	9180
tcagtgaggc	acctatctca	gcgatctgtc	tatttcgttc	atccatagtt	gcctgactcc	9240
ccgtcgtgta	gataactacg	atacgggagg	gcttaaccatc	tggtcccagc	gctgcaatga	9300
taccgcgaga	cccacgctca	ccggctccag	atztatcagc	aataaacccag	ccagccggaa	9360
gggcccagcg	cagaagtggt	cctgcaactt	tatecgcctc	catccagttc	attaattggt	9420
gccgggaagc	tagagtaagt	agttccocag	ttaatagttt	gcgcaacggt	gttgccattg	9480
ctacaggcat	cgtgggtgca	cgctcgtcgt	ttggatggc	ttcattcagc	tccggttccc	9540
aacgatcaag	gcgagttaca	tgatccccca	tgttgtgcaa	aaaagcgggt	agctccttcg	9600
gtcctccgat	cgttgtcaga	agtaagttgg	ccgcagtggt	atcactcatg	gttatggcag	9660
cactgcataa	ttctcttact	gtcatgccat	ccgtaagatg	cttttctgtg	actggtgagt	9720
actcaaccaa	gtcattctga	gaatagtgtg	tgccggcacc	gagttgctct	tgcccggcgt	9780
caacacggga	taataccgcg	ccacatagca	gaactttaa	agtgtctatc	attgaaaaag	9840
acctgcaggg	gggggggggc	gctgaggtct	gctcgtgaa	gaagggttg	ctgactcata	9900
ccaggcctga	atcgccccat	catccagcca	gaaagtgagg	gagccacggt	tgatgagagc	9960
tttgtttag	gtggaccagt	tggtgattht	gaactttgc	tttgccacgg	aacggtctgc	10020
gttgtcggga	agatgcgtga	tctgatcctt	caactcagca	aaagttcgat	ttattcaaca	10080
aagccgccgt	cccgtcaagt	cagcgtaatg	ctctgccagt	gttacaacca	attaaccaat	10140
tctgattaga	aaaactcatc	gagcatcaaa	tgaaactgca	atthattcat	atcaggatta	10200
tcaataccat	atthttgaaa	aagccgtttc	tgtaatgaag	gagaaaaactc	accgagggcag	10260
ttccatagga	tggcaagatc	ctggatccg	tctgcgattc	cgactcgtcc	aacatcaata	10320
caacctatta	atthcccctc	gtcaaaaata	aggttatcaa	gtgagaaatc	accatgagtg	10380
acgactgaat	ccggtgagaa	tggcaaaagc	ttatgcattt	ctttccagac	ttgttcaaca	10440
ggccagccat	tacgctcgtc	atcaaaaatca	ctcgcaccaa	ccaaaaccgtt	attcattcgt	10500
gattgcgcct	gagcggagacg	aaatacgcga	tcgctgttaa	aaggacaatt	acaaacagga	10560
atcgaatgca	accggcgcag	gaacactgcc	agcgcaccaa	caatattttc	acctgaatca	10620
ggatattctt	ctaataacctg	gaatgctgtt	ttcccgggga	tcgcagtggt	gagtaaccat	10680
gcatcatcag	gagtacggat	aaaatgcttg	atggtcggaa	gaggcataaa	ttccgtcagc	10740

-continued

---

cagtttagtc tgaccatctc atctgtaaca tcattggcaa cgetaccttt gccatgttcc 10800  
agaaacaact ctggcgcatac gggcttccca tacaatcgat agattgtcgc acctgattgc 10860  
ccgacattat cgcgagccca tttataccca tataaatcag catccatggt ggaatttaat 10920  
cgcgccctcg agcaagacgt ttccegttga atatggctca taacaccctt tgtattactg 10980  
tttatgtaag cagacagttt tattgttcat gatgatata tttatcttg tgcaatgtaa 11040  
catcagagat tttgagacac aacgtggctt tcccccccc ccctgcaggc caattcggtc 11100  
gatatggcta ttacgaagaa ggtctgtcgc cggagtcctg tgaactttcc caccgcaaca 11160  
gtgaaccgca cggggttgc cggaggccat ttcgttaaaa tgcgcagcca tggtcgttcc 11220  
gtccagcatg gcgtaatact gatcctcgtc ttcggctggc ggatatattg cgatgggctt 11280  
caaaagccgc cgtggtttaa ccagctctac cattccaagg tagcgaactc gaccgcttcg 11340  
aagctcctcc atggtccacg ccgatgaatg acctcggcct tgtaaagacc gttgatcgtc 11400  
tctcgcaggg cgttgctcgtg ctgtcgcgca cgcttccgat agatggctcg atacctgctt 11460  
ctgccaaccg ctccgaatag cgaaggaca cgtattgaac accgcgatcc gagtgatgca 11520  
ctaggccccc atgagcggga cgcgatcat gatgagctc ctgcagggca tcgaggacaa 11580  
agcctgcatg tgctgtcccg ctgcgccccc atccgacaat gcgacgggcg aagacgtcga 11640  
tcacgaaggg cacgtagacg aagccctccc aagtggcgac ataagtacgg acatgcgcaa 11700  
aggctttccc ggtttgtcgc tgatggtgca agagacgctg aagcgcgatc cgatgcgcag 11760  
gcatctgttc gtcttccgcg gtcgtggcgg tggcctgac aaggtcactc gccgaagagc 11820  
tgcatgattg gctcgaacc gagcggggga aattgtcgcg cagttctccc gtcgcgaggg 11880  
cgataaatta catgctcaag cgatgggatg gcattacgtc attcctcgat gacggcccga 11940  
tttgctgac gaacaatgct gccgaacgaa cgctcagagg ctatgtactc gccaggaagt 12000  
catggctggt tgccggatcg gatcgttggt ctgaacgtgc ggcgttcgat gcgacactga 12060  
tcatgagcgc caagctcaat aacatcgatc cgcaggcctg gcttgccgac gtcgcgcccg 12120  
accttgccga cgctccgatc agcaggcttg agcaacagct gccgtggaac tggacatcca 12180  
agacactgag tgetcaggcg gctgaactg cggccttcac cggatactta ccccattatc 12240  
gcagattgcg atgaagcadc agcgtcattc agcaatcttg ccaaagtatg caggctcgcg 12300  
agaatcgacg tgcgaaaccg gctggttgcg ccaaagatcc gcttgccggag cggtcgaaca 12360  
ttcatgctgg gacttcaaga ggtcagtag aggaagaacc ggaaagggtg caccggaaaa 12420  
tatgcgttcc tttggagagc gcctcatgga cgtgaacaaa tcgcccggac caaggatgcc 12480  
acggatacaa aagctcgcga agctcggctc cgtgggtggt ctgtcgtctc gttgtacaac 12540  
gaaatccatt cccattccgc gctcaagatg gcttcccctc ggcagttcat cagggtctaa 12600  
tcaatctagc cgacttgtcc ggtgaaatgg gctgcactcc aacagaaaca atcaacaaa 12660  
catacacagc gacttattca cagcagctca aattacaacg gtatataatc tgccagtcag 12720  
catcatcaca ccaaaagtta ggcccgaata gtttgaaatt agaaagctcg caattgaggt 12780  
ctacaggcca aattcgtctc tagccgtaca atattactca ccggtgcgat gccccccatc 12840  
gtaggtgaag gtgaaatta atgatccatc ttgagaccac agggccacaa cagctaccag 12900  
tttctcaag ggtccaccaa aaacgtaagc gcttacgtac atggtcgata agaaaaggca 12960  
atgttagat gttaacatcc aacgtcgtt tcagggatcg atccaatcgc caaacccctc 13020  
ctccccgcgc gttggccgat tcattaatgc agctggcacg acaggtttcc cgactggaaa 13080

-continued

---

```

gccccgagtg agcgcaacgc aattaatgtg agttagctca ctccattagc accccaggct 13140
ttacacttta tgcttcgggc tcgtatgttg tgtggaattg tgagcggata acaatttcac 13200
acaggaaaca gctatgacca tgattacgcc aagcttgcat gcctgcaggc cgactctaga 13260
ggatctgg                                     13268

```

```

<210> SEQ ID NO 28
<211> LENGTH: 20921
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: vector pKR1223
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: (15083)..(15083)
<223> OTHER INFORMATION: n is a, c, g, or t

```

```

<400> SEQUENCE: 28

```

```

cgcccgagat cctctagagt cgacctgcag gcatgcaagc ttggcgtaat catggtcata 60
gctgtttcct gtgtgaaatt gttatccgct cacaattcca cacaacatac gagccggaag 120
cataaagtgt aaagcctggg gtcgctaatt agtgagctaa ctccattaa ttgcgttgcc 180
ctcaactgcc gctttccagt cgggaaacct gtcgtgccag ctgcattaat gaatcgcca 240
acgcgcgggg agagcggtt tgcgtattgg atcgatccct gaaagcgacg ttggatgta 300
acatctacaa attgcctttt cttatcgacc atgtacgtaa gcgcttacgt ttttggtgga 360
cccttgagga aactggtagc tgtgtgggc ctgtggtctc aagatggatc attaatctcc 420
accttcacct acgatggggg gcatcgcacc ggtgagtaat attgtacggc taagagcgaa 480
tttggcctgt agacctcaat tgcgagcttt ctaatttcaa actattcggg cctaactttt 540
ggtgtgatga tgctgactgg caggatatac accgttgtaa tttgagctcg tgtgaataag 600
tcgctgtgta tgtttgtttg attgtttctg ttggagtgca gccatttca ccggacaagt 660
eggctagatt gatttagccc tgatgaactg ccgaggggaa gccatcttga gcgcggaatg 720
ggaatggatt tcgtgtgaca acgagacgac agaacaccca cgggaccgag cttcgcgagc 780
ttttgtatcc gtggcatcct tggccgggc gatttgttca cgtccatgag gcgctctcca 840
aaggaaacgca tattttccgg tgcaaccttt ccggttcttc ctctactcga cctcttgaag 900
tcccagcatg aatgttcgac cgctccgcaa gcggatcttt ggcgcaacca gccggtttcg 960
cacgtcgatt ctccgagacc tgcatacttt ggcaagattg ctgaatgacg ctgatgcttc 1020
atcgcaatct gcgataatgg ggtaagtatc cggggaagcc cgcaggtcag gccgcctgag 1080
cactcagtgct cttggatgct cagttccacg gcagctgttg ctcaagcctg ctgatcggag 1140
cgtccgcaag gtcggcggcg acgtcggcaa gccaggcctg cggatcgatg ttattgagct 1200
tgggcctcat gatcagtgct gccatgaacg ccgcacgttc agcacaacga tccgatccgg 1260
caaacagcca tgacttctcg ccgagtacat agcctctgag cgttcgttcg gcagcattgt 1320
tcgtcaggca aatcgggccc tcctcaggga atgacgtaat gccatcccat cgcttgagca 1380
tgtaatttat cgcctcggcg acgggagaaac tgccgcgaaa tttccccgc tcggtttcga 1440
gccaatcatg cagctcttcg gcgagtgacc ttgatcaggc caccgccacg accgcggaag 1500
acgaacagat gcctcgccat cggatcgccg ttcagcgtct cttgcacatc cagcgacaaa 1560
ccgggaaagc ctttgcgcat gtcctgactt atgtcgccac ttgggagggc ttcgtctacg 1620
tggcctctgt gatcgacgct ttcgcccgtc gcattgtcgg atggcggggc agccggacag 1680
cacatgcagg ctttgcctc gatgccctcg aggaggtcga tcatgatcgg cgtcccgtc 1740

```

-continued

---

atggcggcct	agtgcacac	tcggatcgcg	gtgttcaata	cgtgtccttt	cgctattccg	1800
agcggttggc	agaagcaggt	atcgagccat	ctatcggaag	cgctcggcag	agcacgacaa	1860
cgccctcgca	gaagcgatca	acggctctta	caaggccgag	gtcattcatc	ggcgtggacc	1920
atggaggagc	ttcgaagcgg	tcgagttcgc	taccttgaa	tggatagact	ggttcaacca	1980
cggcggcttt	tgaagcccat	cggcaatata	ccgccagccg	aagacgagga	tcagtattac	2040
gccatgctgg	acgaagcagc	catggctcgc	cattttaacg	aaatggcctc	cggcaaaccc	2100
ggtgcggttc	acttgttgcg	tgggaaagt	cacgggactc	cgcgcacgag	ccttcttctg	2160
aatagccata	tcgaccgaat	tgacctgcag	gggggggggg	gaaagccacg	ttgtgtctca	2220
aaatctctga	tgttacattg	cacaagataa	aaatatatca	tcatgaacaa	taaaactgtc	2280
tgcttacata	aacagtaata	caaggggtgt	tatgagccat	attcaacggg	aaacgtcttg	2340
ctcagggcgc	cgattaaatt	ccaacatgga	tgctgattta	tatgggtata	aatgggctcg	2400
cgataatgtc	gggcaatcag	gtcgcacaat	ctatcgattg	tatgggaagc	ccgatgcgcc	2460
agagttgttt	ctgaaacatg	gcaaaagtag	cgttgccaat	gatgttacag	atgagatggt	2520
cagactaaac	tggctgacgg	aatttatgcc	tcttcgcgac	atcaagcatt	ttatccgtac	2580
tcctgatgat	gcattggttac	tcaccactgc	gatccccggg	aaaacagcat	tccaggtatt	2640
agaagaatat	cctgattcag	gtgaaaatat	tgttgatgcg	ctggcagtg	tectgcgccg	2700
gttgcatctg	attcctgttt	gtaattgtcc	ttttaacagc	gatcgcgtat	ttcgtctcgc	2760
tcaggcgcga	tcacgaatga	ataacggttt	ggttgatgcg	agtgattttg	atgacgagcg	2820
taatggctgg	cctgttgaa	aagtctggaa	agaaatgcat	aagcttttgc	cattctcacc	2880
ggattcagtc	gtcactcatg	gtgatttctc	acttgataac	cttatttttg	acgaggggaa	2940
attaataggt	tgtattgatg	ttggacgagt	cggaatcgca	gaccgatacc	aggatcttgc	3000
catctatgg	aactgcctcg	gtgagtttct	tccttcatta	cagaaacggc	tttttcaaaa	3060
atatggtatt	gataatcctg	atatgaataa	attgcagttt	catttgatgc	tcgatgagtt	3120
tttctaata	gaattgggta	attggttgta	acactggcag	agcattacgc	tgacttgacg	3180
ggacggcggc	tttgttgaa	aaatcgaa	tttgcagagt	tgaaggatca	gatcacgcat	3240
cttcccgaca	acgcagaccg	ttccgtggca	aagcaaaagt	tcaaaatcac	caactggctc	3300
acctacaaca	aagctctcat	caaccgtggc	tcctcactt	tctggctgga	tgatggggcg	3360
attcaggcct	ggatgatgac	agcaaacact	tcttcacgag	gcagacctca	gcgccccccc	3420
ccccctgcag	gtcttttcca	atgatgagca	cttttaaagt	tctgctatgt	ggcgcgggat	3480
tatccctgtg	tgacgcgggg	caagagcaac	tcggtcgcgc	catacactat	tctcagaatg	3540
acttgggtga	gtactcacca	gtcacagaaa	agcatcttac	ggatggcatg	acagtaagag	3600
aattatgcag	tgctgccata	accatgagtg	ataacactgc	ggccaactta	cttctgacaa	3660
cgatcggagg	accgaaggag	ctaaccgctt	ttttgcacaa	catgggggat	catgtaactc	3720
gccttgatcg	ttgggaaacc	gagctgaatg	aagccatacc	aaacgacgag	cgtgacacca	3780
cgatgcctgt	agcaatggca	acaacgttgc	gcaaacattt	aactggcgaa	ctacttactc	3840
tagcttcccc	gcaacaatta	atagactgga	tggaggcggg	taaagttgca	ggaccacttc	3900
tgcgctcggc	ccttccggct	ggctggttta	ttgctgataa	atctggagcc	ggtgagcgtg	3960
ggtctcgcgg	tatcattgca	gcactggggc	cagatggtaa	gccctcccg	atcgtagtta	4020
tctacacgac	ggggagtcag	gcaactatgg	atgaacgaaa	tagacagatc	gctgagatag	4080

-continued

---

gtgcctcact gattaagcat tgtaactgt cagaccaagt ttactcatat atactttaga	4140
ttgattttaa acttcatttt taatttaaaa ggatctaggt gaagatcctt tttgataatc	4200
tcatgaccaa aatcccttaa cgtgagtttt cgttccactg agcgtcagac cccgtagaaa	4260
agatcaaagg atcttcttga gatccttttt ttctgogcgt aatctgctgc ttgaaaaca	4320
aaaaaccacc gctaccagcg gtggtttgtt tgccggatca agagctacca actccttttc	4380
cgaaggtaac tggttcagc agagcgcaga taccaaatc tgctcttcta gtgtagccgt	4440
agttaggcca ccacttcaag aactctgtag caccgcctac atacctcgtc ctgtaatcc	4500
tgttaccagt ggctgctgcc agtggcgata agtcgtgtct taccgggttg gactcaagac	4560
gatagttacc ggataaggcg cagcggcgg gctgaacggg gggttcgtgc acacagccca	4620
gcttgagcgc aacgacctac accgaactga gatacctaca gcgtgagcta tgagaaagcg	4680
ccacgcttcc cgaagggaga aaggcggaca ggtatccggt aagcggcagg gtcggaacag	4740
gagagcgcac gagggagcct ccagggggaa acgcttggtt tctttatagt cctgtcgggt	4800
ttcggccact ctgacttgag cgtcgatttt tgtgatgctc gtcagggggg cggagcctat	4860
ggaaaaacgc cagcaacgcg gcctttttac ggttctcggc cttttgctgg ccttttgctc	4920
acatgttctt tcctcgttca tccctgatt ctgtggataa ccgtattacc gcctttgagt	4980
gagctgatac cgctcgcgc agccgaacga ccgagcgcag cgagtcagtg agcgaaggaag	5040
cggaagagcg cctgatcggc tattttctcc ttacgcactc gtgcggtatt tcacaccgca	5100
tatggtgcac tctcagtaca atctgctctg atgcgcata gtttaagccag tatacactcc	5160
gctatcgcta cgtgactggg tcatggctgc gcccgcacac ccgccaacac ccgctgacgc	5220
gccctgacgg gcttgtctgc tcccggcacc cgttacaga caagctgtga ccgtctccgg	5280
gagctgcatg tgtcagaggt tttcacogtc atcacgaaa cgcgcgagge aggggtgctt	5340
gatgtggcgc ccggcggctc agtggcgcag gcgcggcttg tccgcgcctt ggtagattgc	5400
ctggccgtag gccagccatt tttgagcggc cagcggcgcg gatagggcca cgcgaagcgg	5460
cggggcgtag ggagcgcagc gaccgaaggg taggcgcttt ttgcagctct tcggctgtgc	5520
gctggccaga cagttatgca caggccagge gggttttaag agttttaata agttttaaag	5580
agttttaggc ggaaaaatcg cctttttctc cttttatata agtcaactac atgtgtgacc	5640
ggttcccaat gtacggcttt gggttcccaa tgtacgggtt ccggttccca atgtacggct	5700
ttgggttccc aatgtacgtg ctatccacag gaaagagacc ttttgcacct tttccctcgt	5760
ctagggcaat ttgccctagc atctgctccg tacattagga accggcggat gcttcgcctc	5820
cgatcaggtt gcggtagcgc atgactagga tcgggcccag ctgccccgc tctccttca	5880
aatcgtactc cggcaggtca tttgaccoga tcagcttgcg cacggtgaaa cagaacttct	5940
tgaactctcc ggcgctgcca ctgcgttcgt agatcgtctt gaacaacat ctggcttctg	6000
ccttgctcgc ggcgcggcgt gccaggcggg agagaaaacg gccgatgccg ggatcgatca	6060
aaaagtaac ggggtgaacc gtcagcacgt ccgggttctt gccttctgtg atctcgcggg	6120
acatccaatc agctagctcg atctcgatgt actccggccg cccggtttcg ctctttacga	6180
tcttgtagcg gctaatcaag gcttcacctc cggataccgt caccaggcgg ccggttcttg	6240
ccttctctgt acgctgcatg gcaacgtgcg tgggttttaa ccgaatgcag gtttctacca	6300
ggctcgtctt ctgctttcgc ccateggctc gccgcagaaa cttgagtacg tccgcaactg	6360
gtggacggaa cacgcggcgc ggttctgtct ccttcccttc ccggtatcgg ttcattggatt	6420
ccggttagatg ggaaccgcgc atcagtaaca ggtcgtaatc ccacacactg gccatgccg	6480

-continued

---

ccggccctgc	ggaaaacctct	acgtgcccgt	ctggaagctc	gtagcggatc	acctcgcag	6540
ctcgtcggtc	acgcttcgac	agacggaaaa	cggccacgtc	catgatgctg	cgactatcgc	6600
gggtgcccac	gtcatagagc	atcggaacga	aaaaatctgg	ttgctcgtcg	cccttgggcg	6660
gcttcctaat	cgaaggcgca	ccggtgccc	gcggttgccg	ggattctttg	cggattcgat	6720
cagcggccgc	ttgccacgat	tcaccggggc	gtgcttctgc	ctcgatgcgt	tgcctgctgg	6780
cggcctgcgc	ggccttcaac	ttctccacca	ggtcatcacc	cagcgcgcgc	cggatttgta	6840
ccgggcccga	tggtttcgca	ccgctcacgc	cgattcctcg	ggcttggggg	ttccagtgcc	6900
attgcagggc	cggcagacaa	cccagccgct	tacgcctggc	caaccgccc	ttcctccaca	6960
catggggcat	tccacggcgt	cggtgccctg	ttgttcttga	ttttccatgc	cgctccttt	7020
agccgctaaa	attcatctac	tcatttattc	atgtgctcat	ttactctggt	agctgcgcga	7080
tgtattcaga	tagcagctcg	gtaatggtct	tgcttgccg	taccgcgtac	atcttcagct	7140
tggtgtgatc	ctccgccggc	aactgaaagt	tgaccgcctt	catggctggc	gtgtctgcc	7200
ggctggccaa	cgttgccagc	ttgctgctgc	gtgcgctcgg	acggccggca	cttagcgtgt	7260
ttgtgctttt	gctcattttc	tctttaoctc	attaactcaa	atgagttttg	atttaatttc	7320
agcggccagc	gcctggacct	cgcgggcagc	gtcgcctcgc	ggttctgatt	caagaacggt	7380
tgtgcggcgc	gcgccagctc	ctgggtagct	cacgcgctgc	gtgatacggg	actcaagaat	7440
gggcagctcg	taccgggcca	gcccctggc	aacctcaccg	ccgatgcgcg	tgctttgat	7500
cggccgcgac	acgacaaagg	ccgcttgtag	ccttccatcc	gtgacctcaa	tgctgctgct	7560
aaccagctcc	accaggctcg	cggtgggcca	tatgtcgtaa	gggcttggtc	gcaccggaat	7620
cagcacgaag	tcggctgctt	tgatcggga	cacagccaag	tccgccctc	ggggcctcc	7680
gtcgatcact	acgaagtcgc	gcccggccgat	ggccttcacg	tcgcggtaaa	tcgtcgggcg	7740
gtcgatgcgc	acaacgggta	gcggttgatc	ttcccgcacg	gcccaccaat	cgcgggcaat	7800
gccctgggga	tcggaatcga	ctaacagaa	atcggccccg	gcgagttgca	gggocggggc	7860
tagatggggt	gcgatggctg	tcttgccctg	cccgcctttc	tggttaagta	cagcagatac	7920
ttcatgcggt	cccttgcgta	ttgttttatt	tactcatcgc	atcatatacg	cagcagaccg	7980
atgacgcaag	ctgttttact	caaatacaca	tcaccttttt	agacggcggc	gctcggtttc	8040
ttcagcggcc	aagctggccg	gccaggccgc	cagcttgcca	tcagacaaac	cggccaggat	8100
ttcatgcagc	cgcacgggtg	agacgtgcgc	ggggcgctcg	aacacgtacc	cggcccgcat	8160
catctccgcc	tcgatctctt	cgtaaatgaa	aaacgggttc	tcctggccgt	cctggtgcgg	8220
tttcatgctt	gttctctctg	gcgttcattc	tcggcgccgc	ccagggcgtc	ggcctcggtc	8280
aatgcgctct	cacggaaggc	accgcgccgc	ctggcctcgc	tgggcgctac	ttctcgcctg	8340
cgctcaagtg	cgcggtacag	ggtcgagcga	tgcaagccaa	gcagtgcagc	cgctctttc	8400
acggtgcggc	cttctcggtc	gatcagctcg	cgggcgtgcg	cgatctgtgc	cggggtgagg	8460
gtaggggcgg	ggccaaaact	cacgcctcgc	gccttgccgc	cctcgcgcc	gctccgggtg	8520
cggctcgatg	ttagggaacg	ctcgaactcg	gcaatgcggg	cgaacacggt	caacaccatg	8580
cggccggccg	gcgtgggtgt	gtcggccca	ggctctgcca	ggctacgcag	gcccgcgccg	8640
gcctcctgga	tgcgctcggc	aatgtccagt	aggtcgcggg	tgctgcgggc	caggcggctc	8700
agcctggtca	ctgtcacaa	gtcggcaggg	cgtagggtgt	caagcatcct	ggccagctcc	8760
ggggcgtcgc	gcctggtgcc	ggtgatcttc	tcggaaaaa	gcttggtgca	gcccgcgcgc	8820

-continued

---

tgcagttcgg	cccgttggtt	ggtcaagtcc	tggtcgtcgg	tgctgacgcg	ggcatagccc	8880
agcaggccag	cggcggcgct	cttgttcatg	gcgtaatgtc	tccggttcta	gtcgcaagta	8940
ttctacttta	tgcgactaaa	acacgcgaca	agaaaacgcc	aggaaaaggg	cagggcggca	9000
gcctgtcgcg	taacttagga	cttgtgcgac	atgtcgtttt	cagaagacgg	ctgcaactgaa	9060
cgtcagaagc	cgactgcact	atagcagcgg	aggggttggg	ccacaggacg	ggtgtggtcg	9120
ccatgatcgc	gtagtcgata	gtggctccaa	gtagcgaagc	gagcaggact	ggcgcgcggc	9180
caaagcggtc	ggacagtgct	ccgagaacgg	gtgcgcatag	aaattgcatc	aacgcatata	9240
gcgctagcag	cacgccatag	tgactggcga	tgctgtcggg	atggacgata	tcccgcaaga	9300
ggcccggcag	taccggcata	accaagccta	tgccctacagc	atccagggtg	acggtgccga	9360
ggatgacgat	gagcgcattg	ttagatttca	tacacggtgc	ctgactgcgt	tagcaattta	9420
actgtgataa	actaccgcat	taaagctagc	ttgcttggtc	gttccgcgctg	aacgtcggct	9480
cgattgtacc	tgcgttcaaa	tactttgcca	tcggttgctg	cgctgcgccg	gtgcgtcggc	9540
tgatctcaog	gatcgactgc	ttctctcgca	acgccatccg	acggatgatg	tttaaaagtc	9600
ccatgtggat	cactccgttg	ccccgtcgtc	caccgtggtg	gggggaaggt	gcacatggct	9660
cagttctcaa	tggaaattat	ctgcctaacc	ggctcagttc	tgcgtagaaa	ccaacatgca	9720
agctccacog	ggtgcaaagc	ggcagcggcg	gcaggatata	ttcaattgta	aatggcttca	9780
tgtccgggaa	atctacatgg	atcagcaatg	agtatgatgg	tcaatatgga	gaaaaagaaa	9840
gagtaattac	caattttttt	tcaattcaaa	aatgtagatg	tccgcagcgt	tattataaaa	9900
tgaaagtaca	ttttgataaa	acgacaaatt	acgatccgtc	gtatttatag	gcgaaagcaa	9960
taaacaaatt	attctaattc	ggaatctttt	atctcgacgt	gtctacattc	acgtccaat	10020
gggggcttag	atgagaaact	tcacgatcga	tgcccttgatt	tcgccattcc	cagataacca	10080
tttcatcttc	agattggtct	gagattatgc	gaaaatatac	actcatatac	ataaactctg	10140
acagtttgag	ctaccaattc	agtgtagccc	attacctcac	ataattcact	caaagtctag	10200
gcagtcctgc	aactcggcgt	caatttgctg	gccactatac	gatagttgctg	caaattttca	10260
aagtcctggc	ctaacatcac	acctctgtcg	gcggcgggtc	ccatttgatg	taaatccacc	10320
atatcgaatt	aattcagact	cctttgcccc	agagatcaca	atggacgact	tcctctatct	10380
ctacgatcta	gtcaggaagt	tcgacgggga	aggtgacgat	accatgttca	ccactgataa	10440
tgagaagatt	agccttttca	atctcagaaa	gaatgctaac	ccacagatgg	ttagagaggc	10500
ttacgcagca	ggtctcatca	agacgatcta	cccagcaat	aatctccagg	agatcaaata	10560
ccttcccaag	aaggttaaag	atgcagtcaa	aagattcagg	actaactgca	tcaagaacac	10620
agagaaagat	atattttctca	agatcagaag	tactattcca	gtatggacga	ttcaaggctt	10680
gcttcacaaa	ccaaggcaag	taatagagat	tggagtctct	aaaaaggtag	ttcccactga	10740
atcaaaggcc	atggagtcaa	agattcaaat	agaggaccta	acagaactcg	ccgtaaagac	10800
tggcgaacag	ttcatacaga	gtctcttacg	actcaatgac	aagaagaaaa	tctctgctca	10860
catggtggag	cacgacacgc	ttgtctactc	caaaaatatac	aaagatacag	tctcagaaga	10920
ccaaagggca	attgagactt	ttcaacaaaag	ggtaatatcc	ggaaacctcc	tcggattcca	10980
tgcccagct	atctgtcact	ttattgtgaa	gatagtgga	aaggaagggtg	gctcctacaa	11040
atgccatcat	tgcgataaag	gaaaggccat	cgttgaagat	gcctctgccc	acagtggctc	11100
caaagatgga	ccccacccca	cgaggagcat	cgtggaaaaa	gaagacgttc	caaccacgtc	11160
ttcaaagcaa	gtggattgat	gtgatatctc	cactgacgta	agggatgacg	cacaatccca	11220

-continued

---

ctatccttgc caagaccctt cctctatata aggaagtcca tttcatttgg agaggacacg 11280  
ctgaaatcac cagtctccaa gcttgcgggg atcgtttcgc atgattgaac aagatggatt 11340  
gcacgcaggt tctccggcgc cttgggtgga gaggtatttc ggctatgact gggcacaaca 11400  
gacaatcggc tgctctgatg ccgcccgtgt ccggtgtgca gcgcaggggc gcccggttct 11460  
ttttgtcaag accgacctgt ccggtgcctt gaatgaactg caggacgagg cagcgcggct 11520  
atcgtagctg gccacgacgg gcgttccttg cgcagctgtg ctgcagcttg tcaactgaagc 11580  
gggaagggac tggtctgatg tgggcgaagt gccggggcag gatctcctgt catctcacct 11640  
tgctcctgcc gagaaagtat ccatcatggc tgatgcaatg cggcggctgc atacgcttga 11700  
tccggctacc tgcccattcg accaccaagc gaaacatcgc atcgagcgag cactgactcg 11760  
gatggaagcc ggtcttctgc atcaggatga tctggacgaa gagcatcagg ggctcgcgcc 11820  
agccgaactg ttcgccaggc tcaaggcgcg catgcccgcg ggcgagatc tcgtcgtgac 11880  
ccatggcgat gcctgcttgc cgaatatcat ggtggaaaat ggccgctttt ctggattcat 11940  
cgactgtggc cggtctgggt tggcggaccg ctatcaggac atagcgttgg ctaccctgga 12000  
tattgctgaa gagcttggcg gcgaatgggc tgaccgcttc ctctgctttt acggtatcgc 12060  
cgctcccgat tcgcagcgca tcgccttcta tcgccttctt gacgagtctt tctgagcggg 12120  
actctggggg tcgaaatgac cgaccaagcg acgcccaccg tgccatcacg agatttcgat 12180  
tccaccgcgc ccttctatga aaggttgggc ttcggaatcg ttttccggga cgcggctggg 12240  
atgatcctcc agcgcgggga tctcatgctg gagttcttcg ccccccgcgg atcgatccaa 12300  
cacttacggt tgcaacgtcc aagagcaaat agaccacgaa cgcgggaagg ttgccgcagc 12360  
gtgtggattg cgtctcaatt ctctcttga ggaatgcaat gatgaatatg atactgacta 12420  
tgaaactttg agggaatact gcctagcacc gtcacctcat aacgtgcacg atgcatgccc 12480  
tgacaacatg gaacatcgtc atttttctga agaattatgc tcgttggagg atgtcgcggc 12540  
aattgcagct attgccaaca tcgaactacc cctcacgcgt gcattcatca atattattca 12600  
tgccgggaaa ggcaagatta atccaactgg caaatcatcc agcgtgattg gtaacttcag 12660  
ttccagcgac ttgattcgtt ttgggtctac ccacgttttc aataaggacg agatggtgga 12720  
gtaaagaagg agtgcgctga agcagatcgt tcaaacattt ggcaataaag tttcttaaga 12780  
ttgaatcctg ttgccggtct tcgcatgatt atcatataat ttctggtgaa ttacgttaag 12840  
catgtaataa ttaacatgta atgcatgacg ttattttatga gatgggtttt tatgattaga 12900  
gtcccgaat tatacattta atacgcgata gaaaacaaaa tatagcgcgc aaactaggat 12960  
aaattatcgc gcgcgggtgc atctatgtta ctagatcgat caaacttcgg tactgtgtaa 13020  
tgacgatgag caatcgagag gctgactaac aaaaggtaaa tcgcgatgga tcgatccatt 13080  
cgccattcag gctgcgcaac tgttgggaag ggcgatcggg gcgggcctct tcgctattac 13140  
gccagctggc gaaaggggga tgtgctgcaa ggcgattaag ttgggtaacg ccagggtttt 13200  
cccagtcacg acgttgtaaa acgacggcca gtgaattcct gcagcccggg ggatccgccc 13260  
actcagggcg cgcctgcgac ggatataatg agccgtaaac aaagatgatt aagtagtaat 13320  
taatacgtac tagtaaaagt ggcaaaagat aacgagaaag aaccaatttc tttgattcgc 13380  
gccttagcgg aaggcatata taagcttga ttattttatt tagtgtaatg atttcgtaca 13440  
accaaagcat ttatttagta ctctcacact tgtgtcgcgg ccggccgcta caggaacagc 13500  
tggtaggggc cctcggcgcg ctctgactgc tccacgatgg tgtagtctc gttgtgggag 13560

-continued

---

gtgatgtcca	gcttggagtc	cacgtagtag	tagccgggca	gctgcacggg	cttcttggcc	13620
atgtagatgg	acttgaactc	caccaggtag	tggccgccgt	ccttcagctt	cagggccttg	13680
tggatctcgc	ccttcagcac	gccgtcgcgg	gggtacaggc	gctcggtgga	ggcctcccag	13740
cccatagtct	tcttctgeat	tacggggccg	tcggagggga	agttcacgcc	gatgaacttc	13800
accttgtaga	tgaaggagcc	gtcctgcagg	gaggagtcc	gggtcacgg	caccacgccg	13860
cgcctctcga	agttcatcac	gcgctcccac	ttgaagccct	cggggaagga	cagcttcttg	13920
tagtcgggga	tgtcggcggg	gtgcttcacg	tacaccttgg	agccgtactg	gaactggggg	13980
gacaggatgt	cccaggcga	gggcaggggg	ccgcccttgg	tcaccttcag	cttggcggtc	14040
tgggtgcctt	cgtagggggg	gccctcgcgc	tcgcctcga	tctcgaactc	gtggcggctt	14100
acggagccct	ccatgcgcac	cttgaagcgc	atgaactcct	tgatgacgtc	ctcggaggag	14160
gccatgggcc	gcttgggggg	ctatggaaga	ctttcttagt	tagttgtgtg	aataagcaat	14220
gttgggagaa	tcgggactac	ttataggata	ggaataaaac	agaaaagtat	taagtgctaa	14280
tgaaatattt	agactgataa	ttaaaatctt	cacgtatgtc	cacttgatat	aaaaacgtca	14340
ggaataaagg	aagtacagta	gaatttaaag	gtactctttt	tatatatacc	cgtgttctct	14400
ttttggctag	ctagttgcat	aaaaaataat	ctatattttt	atcattattt	taaatatctt	14460
atgagatggt	aaatatttat	cataattttt	tttactatta	tttattattt	gtgtgtgtaa	14520
tacatataga	agtttaattac	aaattttatt	tactttttca	ttattttgat	atgattcacc	14580
attaatttag	tgttattatt	tataatagtt	cattttaatc	ttttgtata	tattatgcgt	14640
gcagtaactt	tttctacat	ataactacta	ttacatttta	tttatataat	atttttatta	14700
atgaattttc	gtgataatat	gtaaatattgt	tcattattat	ttcagatttt	ttaaaaaat	14760
ttgtgttatt	atttatgaaa	tatgtaattt	tttttagtatt	tgattttatg	atgataaagt	14820
gttctaaatt	caaaagaagg	gggaaagcgt	aaacattaaa	aaacgctatc	aaacaaaaac	14880
aaaatcttgt	taataaagat	aaaactgttt	gttttgatca	ctgttatttc	gtaataataa	14940
aacattattt	atattttat	tgttgacaac	caaatttgcc	tatcaaatct	aaccaatata	15000
atgcatgcgt	ggcaggtaat	gtactaccat	gaacttaagt	catgacataa	taaaccgtga	15060
atctgaccaa	tgcattgacc	tanctaaatt	gtatttgtga	cacgaagcaa	atgattcaat	15120
tcacaatgga	gatgggaaac	aaataatgaa	gaaccagaa	ctaagaaagc	ttttctgaaa	15180
aataaaataa	aggcaatgtc	aaaagtatac	tgcacatca	gtccagaaag	cacatgatat	15240
ttttttatca	gtatcaatgc	agctagtttt	attttacaat	atcgatatag	ctagtttaaa	15300
tatattgcag	ctagatttat	aaatatttgt	gttattattt	atcatttgtg	taatcctggt	15360
tttagtattt	tagtttatat	atgatgataa	tgtattccaa	atttaaaga	agggaataa	15420
atftaaacaa	gaaaaaaagt	catcaaacaa	aaaacaaatg	aaaggtgga	aagatgttac	15480
catgtaatgt	gaatgttaca	gtatttcttt	tattatagag	ttaacaaatt	aactaatatg	15540
atfttgtaa	taatgataaa	atattttttt	tattattatt	tcataatata	aaaatagttt	15600
acttaatata	aaaaaaatc	tatcgttcac	aacaagttg	gccacctaat	ttaacctatc	15660
atgtaccat	ggaccatatt	aggtaaccat	caaacctgat	gaagagataa	agagatgaag	15720
acttaagtca	taacacaaaa	ccataaaaaa	caaaaataca	atcaaccgtc	aatctgacca	15780
atgcatgaaa	aagctgcaat	agtgagtgcc	gacacaaagc	acatgatftt	cttacaacgg	15840
agataaaacc	aaaaaaat	ttcatgaaca	acctagaaca	aataaagctt	ttatataata	15900
aatatataaa	taataaagg	ctatggaata	atatacttca	atattttgg	atnaataaa	15960

-continued

---

ttggtggcgg ggttgatata ttatatacaca cctaaagtca cttcaatctc attttcactt 16020  
 aacttttatt ttttttttct ttttatttat cataaagaga atattgataa tatacttttt 16080  
 aacatatttt tatgacattt tttattggtg aaaacttatt aaaaatcata aattttgtaa 16140  
 gttagattta tttaaagagt tctcttctt attttaaatt ttttaataaa tttttaaata 16200  
 actaaaatt gtgttaaaaa tgtaaaaaa gtgtgttatt aacccttctc ttcgaggatc 16260  
 cgtaccgagc tcggatccac tagtaacggc cgccagtgtg ctggaattca ggtcctgcag 16320  
 gtctactctt tacatgttct ttactcctc tcaaaatttc cttttttgt tggctctctc 16380  
 cgaacgagtt ggagaaatcg ttaaccctaa tcgaagatct agattcctct acatacgttt 16440  
 gatctctctc tcagtatgga ttcaaaagcg ccaaggagat actactcaca cggagtgtgt 16500  
 gcgagacagc aagatttcgc aacagatata gttacgagaa gaagacctta tgtcccttac 16560  
 gaccgtccaa ataagtttc aaggagtctg gtttggacgt caaaagagta caaatcacc 16620  
 gagggcaata atagcccaag gaccaatgat gtgtcaccga aaccaccagt tttaggtttg 16680  
 gcgaggaaga atgctgcttg tgggccaatg agatcttcta gtctcagaaa atgggtatgt 16740  
 aagtattgga aagatggaaa gtgcaagagg ggtgagcagt gccagttctt aactcttgg 16800  
 tcttgtttcc ctggtatggc catggtagct tctcttgaag ggcacaataa ggaactaaag 16860  
 gggatcgctc tcctgaggg ttcaagataa ctcttttcag tcagtattga tggtagattg 16920  
 cgagtttggg actgcaattc tggtcagtgt gtacattcca tcaacctga cgcagaagca 16980  
 gggctctctaa tcagtgaagg cccttgggtt ttcccttggt tgccaaacgc tataaaggct 17040  
 tttaacgttc aaaccagtc agatttgcac ctcaagcag caggggtggt tggtcaggtg 17100  
 aatgcaatga ctattgaaa cggaaatgct tttgctggaa caagtctctg tagtatctta 17160  
 gtctggaaaag ctactacaga ctctgagct gatccattca aatacttgac atctcttgag 17220  
 ggacatagtg gtgaagtcac ttgttttct gttggaggtc aaatgctata ctctggtct 17280  
 gtcgataaaa caatcaagat gtgggatctc aacacctgc aatgtataat gacctgaag 17340  
 caacataccg gcactgtcac ttcaactctta tggttggata aatgtttgat atcgtcttc 17400  
 ttggatggga ccataaaagt ttgggttct tctgaaaacg gaactctgaa agttgttcaa 17460  
 actcgcagac aagaacagag tagtgttcat gctcttctg gtatgcatga tgcagaagcc 17520  
 aaaccgataa tattctgctc ttacaaaaac ggaaccgttg gcattttcga cctaccatct 17580  
 tttcaagaaa gaggaaggat gttctctacg cacacgatcg ccacactcac aattggtcct 17640  
 caaggattgt tattcagtgg agacgagagt ggtaacttgc gtgtatggac cttagctgct 17700  
 ggcaacaaaag tttagtctt tcgactaaag aattctgatt taattttgtg gtttatatgt 17760  
 tgagttaact gttaagagag ttttattttg taataggtgt atcagtcaat aaacaatctt 17820  
 tgtatcaacc aaatgtaatt tttctctgta attcgatttc agagttttta cttaagata 17880  
 aacaaactct ttcacacatc atttaatgaa agtgagagaag cttaaaaaac aaacaaagaa 17940  
 actgatccat ttttggcggg tcttcttcta ctcttattca tatgtgttaa cgaactatag 18000  
 cgtaaaatc agagcaagcg atctccgatt tgaacgtggc taccaccgga ggcccaccac 18060  
 tacggcgat acgctctaag tgaggattaa agtgctctgg tggtagcgt gaagaaactc 18120  
 gccatgggt tttgttatct ctgcagccaa gtgctgtctc ttcttcgcca cttctcatca 18180  
 agctacagtg aatttaaaaa tggcgtctt ctttgatctc gtatacataa gctggattgg 18240  
 tttcttaaac aaattcctc ccttttgggt cttctgggtt tgcctgttaa gtgtttgtgt 18300

-continued

---

ttttgcctct gagaaaaaat cgcggccgca tggagagatc tcaacggcag tctcctccgc	18360
caccgtcgcc gtcctcctcc tcgtcctccg tctccgcgga caccgtcctc gtcctcctccg	18420
gaaagaggcg gagggcgcg acggccaagg cgggcgcca gcctaataag aggatccgca	18480
aggacccccg cgcgcccgcc gcggggaaga ggagctccgt ctacagggga gtcaccaggc	18540
acaggtggac gggcaggttc gaggcgcac tctgggacaa gcaactgcctc gccgcgctcc	18600
acaacaagaa gaaaggcagg caagtctacc tggggcgcta tgacagcgag gaggcagctg	18660
ctcgtgccta tgacctcgca gctctcaagt actggggctc tgagactctg ctcaacttcc	18720
ctgtggagga ttactccagc gagatgccgg agatggaggc cgtgtcccgg gaggagtacc	18780
tggctcctcc ccgcccgagg agcagcggct tctccagggg cgtctccaag tacagaggcg	18840
tcgccaggca tcaccacaac gggaggtggg aggcacggat tgggcgagtc tttgggaaca	18900
agtaccteta cttgggaaca tttgacactc aagaagaggc agccaaggcc tatgaccttg	18960
cggccattga atacctgggc gtaaatgctg taaccaactt cgacatcagc tgctacctgg	19020
accaccctct gttcctggca cagctccaac aggagccaca ggtggtgccg gcaactcaacc	19080
aagaacctca acctgatcag agcgaaaccc gaactacaga gcaagagccg gactcaagcg	19140
aagccaagac accggatggc agtgcagAAC ccgatgagaa cgcggtgctc gacgacaccg	19200
cggagcccct caccacagtc gacgacagca tcgaagaggg cttgtggagc ccttgcattg	19260
attacgagct agacacctg tcgagaccaa actttggcag ctcaatcaat ctgagcagct	19320
ggttcctgta cgcgacttc gactgcaaca tcggatgcct gttcagatggg tgttctcggg	19380
ctgacgaagg aagcaaggat ggtgtaggtc tggcagattt cagtctgttt gaggcaggtg	19440
atgtccagct gaaggatggt ctttcggata tggaaagagg gatacaacct ccagcgatga	19500
tcagtgtgtg caacgcggcc gcaagtatga actaaaatgc atgtagggtg aagagctcat	19560
ggagagcatg gaatattgta tccgacctg taacagtata ataactgagc tccatctcac	19620
ttctctatg aataaacaac gtagttatg atatattaac actctatcta tgcaccttat	19680
tgttctatga taaatttctc cttattatta taaatcatct gaatcgtgac ggcttatgga	19740
atgcttcaaa tagtacaana acaaatgtgt actataagac tttctaaca attctaacct	19800
tagcattgtg aacgagacat aagtgttaag aagacataac aattataatg gaagaagttt	19860
gtctccattt atatattata tattaccac ttatgtatta tattaggatg ttaaggagac	19920
ataacaatta taaagagaga agtttgatc catttatata ttatatacta cccatttata	19980
tattatactt atccacttat ttaatgtctt tataaggttt gatccatgat atttctaata	20040
ttttagttga tatgtatag aaaggttact atttgaactc tcttactctg tataaagggt	20100
ggatcatcct taaagtgggt ctatttaatt ttattgcttc ttacagataa aaaaaaatt	20160
atgagttggg ttgataaatt attgaaggat ttaaaataat aataaataac atataatata	20220
tgtatataaa tttattataa tataacattt atctataaaa aagtaaatat tgtcataaat	20280
ctatacaatc gtttagcctt gctggacgaa tctcaattat ttaaagcaga gtaaacatat	20340
ttgacttttt ggttatttaa caaattatta ttaaacacta tatgaaattt tttttttat	20400
cagcaagaa taaaattaaa ttaagaagga caatgggtgc ccaatcctta tacaaccaac	20460
ttccacaaga aagtcaagtc agagacaaca aaaaaacaag caaaggaaat tttttaattt	20520
gagttgtctt gtttctgca taatttatgc agtaaaacac tacacataac ctttttagca	20580
gtagagcaat ggttgaccgt gtgcttagct tcttttattt tattttttta tcagcaagaa	20640
ataaataaaa taaatgaga cacttcaggg atgtttcaac aagctctaga gggccaatt	20700

-continued

---

```

cgccctatag tgagtcgtat tacaattcac tggccgctcgt tttacaacgt cgtgactggg 20760
aaaaccctgg cgttacccaa cttaatcgcc ttgcagcaca tcccccttc gccagctggc 20820
gtaatagcga agaggcccgcc accgatcgcc cttcccaaca gttgcgcagc ctatacgtac 20880
gagatccggc cggccagatc ctgcaggaga tccaagcttg g 20921

```

```

<210> SEQ ID NO 29
<211> LENGTH: 4906
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: vector pKR268

```

```

<400> SEQUENCE: 29

```

```

ggccgcatga gccgtaaagg ttcaatacaa cgagtgcttg ttttcttagg gacaagcatt 60
gtacttatgt atgattctgt gtaaccatga gtcctccacg ttgtactaat gtgaagggca 120
aaaataaaac acagaacaag ttcggttttc tcaataaatg tgaaggtaga aaatggaacc 180
atgcctctct tcttgcatgt gatttaaaat attagcagat ggtaccgtac gtgggcgcat 240
cccccgggct gcaggaattc actggccgctc gttttacaac gtcgtgactg ggaaaaccct 300
ggcgttacc aacttaatcg ccttgccagca catccccctt tcgccagctg gcgtaatagc 360
gaagaggccc gcaccgatcg ccttcccaa cagttgcgca gcctgaatgg cgaatggcgc 420
ctgatgcggt attttctcct tacgcactcg tgcggtatth cacaccgat atggtgcaact 480
ctcagtacaa tctgctctga tgcgcatag ttaagccagc cccgacacc gccaacacc 540
gctgacgcgc cctgacgggc ttgtctgctc ccggcatccg cttacagaca agctgtgacc 600
gtctccggga gctgcatgtg tcagaggttt tcaccgcat caccgaaacg cgcgagacga 660
aagggcctcg tgatacgctc atttttatag gttaatgtca tgataaat ggtttcttag 720
acgtcagggtg gcacttttcc gggaaatgtg cgcggaacc ctatttgtht attttctaa 780
atacatcoa atatgtatcc gctcatgaga caataaccct gataaatgct tcaataatat 840
tgaaaaagga agagtatgag tattcaacat ttccgtgctc cccttatcc ctttttgcg 900
gcattttgcc ttctgtttt tgctcaccga gaaacgctgg tgaaagtaa agatgctgaa 960
gatcagttgg gtgcacgagt gggttacatc gaactggatc tcaacagcgg taagatcctt 1020
gagagtthtc gccccgaaga acgttttcca atgatgagca cttttaaagt tctgctatgt 1080
ggcgcggtat tatccgctat tgacgcccgg caagagcaac tcggtcgccg catacactat 1140
tctcagaatg acttggttga gtactcacca gtcacagaaa agcatcttac ggatggcatg 1200
acagtaagag aattatgcag tgctgccata accatgagtg ataacactgc ggccaactta 1260
cttctgacaa cgatcggagg accgaaggag ctaaccgctt ttttgacaa catgggggat 1320
catgtaactc gccttgatcg ttggaaaccg gagctgaatg aagccatacc aaacgacgag 1380
cgtgacacca cgatgcctgt agcaatggca acaacgcttc gcaaactatt aactggcgaa 1440
ctacttactc tagcttcccg gcaacaatta atagactgga tggaggcgga taaagttgca 1500
ggaccacttc tgcgctcggc ccttccggtt ggctggthta ttgctgataa atctggagcc 1560
ggtgagcgtg ggtctcggg tatcattgca gcaactgggc cagatggtaa gccctccgct 1620
atcgtagtta tctacacgac ggggagtcag gcaactatgg atgaacgaaa tagacagatc 1680
gctgagatag gtgcctcact gattaagcat tggtaactgt cagaccaagt ttactcatat 1740
atactttaga ttgatttaaa acttcatttt taatttaaaa ggatctaggt gaagatcctt 1800

```

-continued

---

tttgataatc tcatgaccaa aatcccttaa cgtgagtttt cgttccactg agcgtcagac	1860
cccgtagaaa agatcaaagg atcttcttga gatccttttt ttctgcgcgt aatctgctgc	1920
ttgcaacaaa aaaaaccacc gctaccagcg gtggtttggt tgccggatca agagctacca	1980
actcttttcc cgaaggtaac tggtctcagc agagcgcaga taccaaatac tgtccttcta	2040
gtgtagccgt agttaggcca ccacttcaag aactctgtag caccgcctac atacctcgct	2100
ctgctaatcc tgttaccagt ggtgctgcc agtgcgata agtcgtgtct taccgggttg	2160
gactcaagac gatagttacc ggataaggcg cagcggtcgg gctgaacggg gggttcgtgc	2220
acacagccca gcttggagcg aacgacctac accgaactga gatacctaca gcgtgagcta	2280
tgagaaagcg ccacgcttcc cgaagggaga aaggcggaca ggatccggg aagcggcagg	2340
gtcggaacag gagagcgcac gagggagctt ccaggggaa acgcctggtt tctttatagt	2400
cctgtcgggt ttcgccacct ctgacttgag cgtcgatfff tgtgatgctc gtcagggggg	2460
cggagcctat ggaaaaacgc cagcaacgcg gcctttttac ggttcctggc cttttgctgg	2520
ccttttgctc acatgttctt tctgcggtt tcccctgatt ctgtggataa ccgtattacc	2580
gcctttgagt gagctgatac cgtcgcgccg agccgaacga ccgagcgcag cgagtcagtg	2640
agcaggaag cggaaagagc cccaatacgc aaaccgcctc tcccgcgcg ttggccgatt	2700
cattaatgca gctggcacga caggtttccc gactggaaag cgggcagtga gcgcaacgca	2760
attaatgtga gttagctcac tcattaggca cccaggctt tacactttat gcttccggct	2820
cgtatgttgt gtggaattgt gagcggataa caatttcaca caggaaacag ctatgacct	2880
gattacgcca agcttgcctg cctgcaggtc gactcgacgt acgatcccac atgcaagttt	2940
ttatttcaat cccttttctt ttgaataact gaccaagaac aacaagaaaa aaaaaaaaa	3000
agaaaaggat ctttttgaaa ggatattttt cgctcctatt caaatactgt atttttacca	3060
aaaaaactgt atttttccta cactctcaag ctttgttttt cgcttcgact ctcatgattt	3120
ccttcatacg ccaatcactc tatttataaa tggcataagg tagtgtgaac aattgcaaa	3180
cttgtcatca aaagcttgca atgtacaaat taatgttttt catgcttttc aaaattatct	3240
gcacccccta gctattaatc taacatctaa gtaaggctag tgaatttttt cgaatagtca	3300
tgcagtgcac taatttcccc gtgactatft tggctttgac tccaacactg gccccgtaca	3360
tccgtccctc attacatgaa aagaaatatt gtttatatcc ttaattaaaa atattgtccc	3420
ttctaaatft tcatatagtt aattattata ttactttttt ctctattcta ttagttctat	3480
tttcaaatft ttttttatgc atatgtaaag tacattatat ttttgctata tacttaata	3540
tttctaaatft attaaaaaaa gactgatatg aaaaatttat tctttttaaa gctatatcat	3600
tttatatata ctttttcttt tcttttcttt cttttctat tcaatttaat aagaaataaa	3660
ttttgtaaat tttttattat caatttataa aaatatttta ctttatatgt tttttcacat	3720
ttttgttaaa caaatcatat cattatgatt gaaagagagg aaattgacag tgagtaataa	3780
gtgatgagaa aaaaatgtgt tatttcctaa aaaaaacctc aacaacatg tatctactct	3840
ctatttcatc tatctctcat ttcatttttc tctttatctc tttctttatt tttttatcat	3900
atcatttccac attaatattt tttactctct ttattttttc tctctatccc tctcttattt	3960
ccactcatat atacactcca aaattggggc atgcctttat cactactcta tctctccac	4020
taaatcattt aaatgaaact gaaaagcatt ggcaagtctc ctcccctcct caagtgattt	4080
ccaactcagc attggcatct aattgattca gtatatctat tgcagtgtga aaagtctttc	4140
cacaatacat aactattaat taactttaa taaataaagg ataaaaatatt ttttttctt	4200

-continued

---

```

cataaaatta aaatatgtta ttttttgttt agatgtatat tcgaataaat ctaaataatat 4260
gataatgatt ttttatattg attaaacata taatcaatat taaatatgat atttttttat 4320
ataggttgta cacataatth tataaggata aaaaatatga taaaaataaa ttttaaatat 4380
ttttatattt acgagaaaaa aaaatattht agccataaat aaatgaccag catatthttac 4440
aaccttagta attcataaat tcctatatgt atatttgaaa ttaaaaacag ataatcgтта 4500
agggaaaggaa tcctacgtca tctcttgcca tttgtttttc atgcaaacag aaagggacga 4560
aaaaccacct caccatgaat cactcttcac accatthtta ctagcaaaaca agtctcaaca 4620
actgaagcca gctctctttc cgtttctttt tacaacactt tctttgaaat agtagtattt 4680
ttttttcaca tgatttatta acgtgcaaaa agatgcttat tgaatagagt gcacatttgt 4740
aatgtactac taattagaac atgaaaaagc attgttctaa cacgataatc ctgtgaaggc 4800
gttaactcca aagatccaat ttcactatat aaattgtgac gaaagcaaaa tgaattcaca 4860
tagctgagag aaaaaggaaa ggtaactaa gaagcaatac ttcagc 4906

```

```

<210> SEQ ID NO 30
<211> LENGTH: 10528
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: vector pKR1143
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: (7136)..(7136)
<223> OTHER INFORMATION: n is a, c, g, or t

```

```

<400> SEQUENCE: 30

```

```

gtacgagatc cggccggcca gatcctgcag gagatccaag cttggcgcgc cgttctatag 60
tgtcacctaa atcgatatgt tatgatcat aaggttatgt attaattgta gccgcgttct 120
aacgacaata tgtccatatt gtgcactctc agtacaatct gctctgatgc cgcattagta 180
agccagcccc gacaccgcc aacaccgct gacgcgcct gacgggcttg tctgctcccg 240
gcattccgctt acagacaagc tgtgaccgtc tccgggagct gcattgttca gaggthttca 300
ccgtcatcac cgaaacgcgc gagacgaaag ggcctcgtga tacgcctatt tttataggth 360
aatgtcatga ccaaaatccc ttaactgag ttttcgttcc actgagcgtc agaccctgta 420
gaaaagatca aagatcttc ttgagatcct tttttctgc gcgtaactctg ctgcttgcaa 480
acaaaaaac caccgctacc agcgggtggt tgtttgccgg atcaagagct accaactctt 540
tttccgaagg taactggctt cagcagagcg cagataccaa atactgtcct tctagtgtag 600
ccgtagttag gccaccactt caagaactct gtaccaccgc ctacatacct cgtctgtgta 660
atcctgttac cagtggctgc tgcactggc gataagctgt gtcttaccgg gttggactca 720
agacgatagt taccggataa ggcgcagcgg tcgggctgaa cggggggttc gtgcacacag 780
cccagcttgg agcgaacgac ctacaccgaa ctgagatacc tacagcgtga gcattgagaa 840
agcgcacgc tccccgaagg gagaaaggcg gacaggatc cggtaagcgg cagggtcggaa 900
acaggagagc gcacgaggga gcttccaggg ggaaacgctt ggtatcttta tagtctgtc 960
gggtttcggc acctctgact tgagcgtcga tttttgtgat gctcgtcagg gggggcggagc 1020
ctatggaaaa acgccagcaa cgcggccttt ttaccggttcc tggccttttg ctggcctttt 1080
gtcacatgt tctttctgc gttatcccct gattctgtgg ataaccgtat taccgccttt 1140
gagtgagctg ataccgctcg ccgcagccga acgaccgagc gcagcagctc agtgagcag 1200

```

-continued

---

gaagcggaaag agcgcccaat acgcaaaccg cctctccccg cgcgttgccc gattcattaa	1260
tgcaggttga tcgattcgac atcgatctag taacatagat gacaccgcgc gcgataattt	1320
atcctagttt gcgcgctata ttttgttttc tatecggtat taaatgtata attgcgggac	1380
tctaatacata aaaaccatc tcataaataa cgtcatgcat tacatgttaa ttattacatg	1440
cttaacgtaa ttcaacagaa attatatgat aatcatcgca agaccggcaa caggattcaa	1500
tcttaagaaa ctttattgcc aaatgtttga acgatctgct tgcagcact ccttctttag	1560
gtacctcaact attcctttgc cctcggacga gtgctggggc gtcggtttcc actatcggcg	1620
agtacttcta cacagccatc ggtccagacg gccgcgcttc tgcgggcgat ttgtgtacgc	1680
ccgacagtcc cggtccgga tccgacgatt gcgtcgatc gacctgcgc ccaagctgca	1740
tcatcgaat tgcctcaac caagctctga tagagttggt caagaccaat gcggagcata	1800
tacgcccgga gccgcggoga tctctcaagc tccggatgcc tccgctcga gtagcgcgc	1860
tgtgctcca tacaagccaa ccacggctc cagaagaaga tgttggcgac ctctgattgg	1920
gaatccccga acatcgctc gctccagtc atgaccgctg ttatgcggcc attgtccgctc	1980
aggacattgt tggagccgaa atcccgctgc acgaggtgcc ggacttcggg gcagtcctcg	2040
gccaaagca tcagctcacc gagagcctgc gcgacggacg cactgacggt gtcgtccatc	2100
acagtttgcc agtgatacac atggggatca gcaatcgccg atatgaaatc acgccatgta	2160
gtgtattgac cgattccttg cgtccgaat gggccgaacc cgctcgtctg gctaagatcg	2220
gccgcagcga tcgcatccat ggctccgcg accggctgca gaacagcggg cagttcgggt	2280
tcaggcaggt cttgcaacgt gacaccctgt gcacggcggg agatgcaata ggtcaggctc	2340
tcgctgaatt cccaatgct aagcaattcc ggaatcgga gcgcggccga tgcaaagtgc	2400
cgataaacat aacgatcttt gtgaaacca tcggcgcagc tatttaccg caggacatat	2460
ccacgccctc ctacatcgaa gctgaaagca cgagattctt cgcctccga gagctgcac	2520
aggcgggaga cgctgtcgaa cttttcgatc agaaactctc cgacagacgt cgcggtgagt	2580
tcaggccttt tcattggtta ataagaagag aaaagagttc ttttgttatg gctgaagtaa	2640
tagagaaatg agctcgagcg tgcctctcc aaatgaaatg aacttcctta tatagaggaa	2700
gggtcttgcg aaggatagtg ggattgtgct tcatccctta cgctcagtgga gatgtcacat	2760
caatccactt gctttgaaga cgtggttga acgtctctt tttccacgat gctcctcgtg	2820
ggtgggggtc catctttggg accactgtcg gcagaggcat cttgaatgat agcctttcct	2880
ttatcgcaat gatggcattt gtaggagcca ccttctcttt ctactgtcct ttctgtagaag	2940
tgacagatag ctgggcaatg gaatccgagg aggtttcccg aaattatcct ttgttgaaaa	3000
gtctcaatag ccctttggtc ttctgagact gtatctttga catttttga gtagaccaga	3060
gtgtcgtgct ccaccatggt gacgaagatt ttcttctgt cattgagtcg taaaagactc	3120
tgtatgaaat gttcgcagc cttcacggcg agttctgtta gatcctcgat ttgaatctta	3180
gactccatgc atggccttag attcagtagg aactacctt ttagagactc caatctctat	3240
tacttgccctt ggtttatgaa gcaagccttg aatcgtccat actggaatag tactttctgat	3300
cttgaaaaat atgtctttct ctgtgttctt gatgcaatta gtctgaatc ttttgactgc	3360
atctttaacc ttcttgggaa ggtatttgat ctctgggaga ttgttactcg ggttagatcgt	3420
cttgatgaga cctgctcgt aggcctctc aaccatctgt gggtcagcat tctttctgaa	3480
attgaagagg ctaacctctc cattatcagt ggtgaacata gtgtcgtcac cttcaccttc	3540
gaacttcctt cctagatcgt aaagatagag gaaatcgtcc attgtaatct ccggggcaaa	3600

-continued

---

ggagatctct	tttggggctg	gatcaactgct	gggccttttg	gttcctagcg	tgagccagtg	3660
ggctttttgc	tttgggtggc	ttgttagggc	cttagcaaaag	ctcttgggct	tgagttgagc	3720
ttctcctttg	gggatgaagt	tcaacctgtc	tgtttgctga	cttgttgtgt	acgcgctcagc	3780
tgctgctctt	gcctctgtaa	tagtggcaaa	tttcttgtgt	gcaactccgg	gaacgcggtt	3840
tgttgccgcc	ttgtacaac	cccagtcate	gtatataccg	gcatgtggac	cgttatacac	3900
aacgtagtag	ttgatatgag	ggtgttgaat	acccgattct	gctctgagag	gagcaactgt	3960
gctgttaagc	tcagattttt	gtgggattgg	aattggatcg	atctcgatcc	cgcgaaatta	4020
atacactca	ctatagggag	accacaacgg	tttccctcta	gaaataattt	tgtttaactt	4080
taagaaggag	atatacccat	ggaaaagcct	gaactcaccg	cgacgtctgt	cgagaagtgt	4140
ctgatcgaaa	agttcgacag	cgtctccgac	ctgatgcagc	tctcggaggg	cgaagaatct	4200
cgtgctttca	gcttcgatgt	aggagggcgt	ggatatgtcc	tgcgggtaaa	tagctgcgcc	4260
gatggtttct	acaagatcg	ttatgtttat	cggcactttg	catcggccgc	gctcccatt	4320
cgggaagtgc	ttgacattgg	ggaattcagc	gagagcctga	cctattgcat	ctcccgcgt	4380
gcacagggtg	tcacgttgca	agacctgcct	gaaaccgaac	tgcccgtgt	tctgcagccg	4440
gtcgcggagg	ctatggatgc	gatcgtgcg	gccgatctta	gccagacgag	cgggttcggc	4500
ccattcggac	cgcaaggaat	cgtcaatac	actacatggc	gtgatttcat	atgcgcgatt	4560
gctgatcccc	atgtgatca	ctggcaaac	gtgatggacg	acaccgtcag	tgcgtccgtc	4620
gcgcaggctc	tcgatgagct	gatgctttgg	gccgaggact	gccccgaagt	ccggcacctc	4680
gtgcacgcgg	atttcggctc	caacaatgtc	ctgacggaca	atggccgcat	aacagcggtc	4740
attgactgga	gcgagggcat	gttcggggat	tccaataacg	aggtcgccaa	catcttcttc	4800
tggaggccgt	ggttggcttg	tatggagcag	cagacgcgct	acttcgagcg	gaggcatccg	4860
gagcttgca	gatcgcgcg	gctccggggc	tatatgtctc	gcattggctt	tgaccaactc	4920
tatcagagct	tggttgacgg	caatttcgat	gatgcagctt	gggcgcaggg	tcgatgcgac	4980
gcaatcgtcc	gatccggagc	cgggactgtc	gggcgtacac	aaatcgcccc	cagaagcgcg	5040
gccgtctgga	ccgatggctg	tgtagaagta	ctcgcgcgata	gtggaaaacc	acgccccagc	5100
actcgtccga	gggcaaagga	atagttaggt	acagcttggg	tcgatccggc	tgctaacaaa	5160
gccccaaaag	aagctgagtt	ggctgctgcc	accgctgagc	aataactagc	ataaccctt	5220
ggggcctcta	aacgggtctt	gaggggtttt	ttgctgaaag	gaggaactat	atccggatga	5280
tcgggcgcgc	cgtcgacgga	tccactagtt	ctagagcggc	ccgcgcgctc	gacggatata	5340
atgagccgta	aacaaagatg	attaagtagt	aattaatac	tactagtaaa	agtggcaaaa	5400
gataacgaga	aagaaccaat	ttctttgcat	tcggccttag	cgggaaggcat	atataagctt	5460
tgattatfff	atftagtgt	atgatttctg	acaaccaaa	catftattta	gtactctcac	5520
acttgtgtcg	cggccggccg	ctacaggaac	aggtggtggc	ggccctcggc	gcgctcgtac	5580
tgctccacga	tggtgtagtc	ctcgttgtgg	gaggtgatgt	ccagcttggg	gtccacgtag	5640
tagtagccgg	gcagctgcac	gggcttcttg	gccatgtaga	tggacttggg	ctccaccagg	5700
tagtggccgc	cgtccttcag	cttcagggcc	ttgtggatct	cgcccttcag	cacgcctcgc	5760
cgggggtaca	ggcgcctcgg	ggaggcctcc	cagcccatag	tcttctctcg	cattacgggg	5820
ccgtcggagg	ggaagtccac	gccgatgaac	ttcaccttgt	agatgaagga	gccgtcctgc	5880
agggaggagt	cctgggtccac	ggtcaccacg	ccgcctcctc	cgaagtccat	cacgcctcgc	5940

-continued

---

cacttgaagc cctcggggaa ggacagcttc ttgtagtcgg ggatgtcggc ggggtgcttc	6000
acgtacacct tggagccgta ctggaactgg ggggacagga tgtcccaggc gaagggcagc	6060
gggcgcacct tggtcacctt cagcttggcg gtctgggtgc cctcgtaggg gggccctcg	6120
ccctcgcctt cgatctcgaa ctctgtggcg ttcacggagc cctccatgcg caccttgaag	6180
cgcatgaact ccttgatgac gtctcggag gaggccatgg gccgcttggg gggctatgga	6240
agactttctt agttagtgtg gtgaataagc aatgttggga gaatcgggac tacttatagc	6300
ataggaataa aacagaaaag tattaagtgc taatgaaata tttagactga taattaaat	6360
cttcacgtat gtccacttga tataaaaacg tcaggaataa aggaagtaca gtagaattta	6420
aaggctactt ttttatatat acccgtgttc tctttttggc tagctagtgg cataaaaaat	6480
aatctataatt tttatcatta ttttaaatat cttatgagat ggtaaataatt taccataatt	6540
ttttttacta ttattttatta tttgtgtgtg taatacatat agaagttaat tacaattttt	6600
atttactttt tcattatttt gatatgattc accattaatt tagtgttatt atttataata	6660
gttcatttta atctttttgt atatatatg cgtgcagtac ttttttccta catataacta	6720
ctattacatt ttattttatat aatattttta ttaatgaatt ttcgtgataa tatgtaatat	6780
tgttcattat tatttcagat tttttaaaaa tatttgtgtt attatttatg aaatagttaa	6840
tttttttagt atttgatttt atgatgataa agtgttctaa attcaaaaga agggggaaaag	6900
cgtaaacatt aaaaaacgtc atcaaaacaaa acaaaaatct tgtaataaaa gataaaactg	6960
tttgttttga tcaactgttat ttcgtaatat aaaaacatta tttatattta tattgttgac	7020
aaccaaattt gcctatcaaa tctaaccaat ataatgcatg cgtggcaggc aatgtactac	7080
catgaactta agtcatgaca taataaacgg tgaatctgac caatgcatgt acctanctaa	7140
attgtatttg tgacacgaag caaatgattc aattcacaat ggagatggga aacaaataat	7200
gaagaacca gaactaagaa agcttttctg aaaaataaaa taaaggcaat gtcaaaagta	7260
tactgcatca tcagtccaga aagcacatga tattttttta tcagtatcaa tgcagctagt	7320
tttattttac aatatcgata tagctagtgt aaatatattg cagctagatt tataaatatt	7380
tgtgttatta tttatcattt gtgtaatcct gtttttagta ttttagttta tatatgatga	7440
taatgtattc caaatttaaa agaagggaaa taaattttaa caagaaaaaa agtcatcaaa	7500
caaaaaacaa atgaaagggt ggaaagatgt taccatgtaa tgtgaatgtt acagtatttc	7560
ttttattata gagttaacaa attaactaat atgattttgt taataatgat aaaatatttt	7620
ttttattatt atttcataat ataaaaatag tttacttaat ataaaaaaaa ttctatcgtt	7680
cacaacaaaag ttggccacct aatttaacca tgcattgtacc catggaccat attaggtaac	7740
catcaaacct gatgaagaga taaagagatg aagacttaag tcataacaca aaaccataaa	7800
aaacaaaaat acaatcaacc gtcaatctga ccaatgcatg aaaaagctgc aatagtgagt	7860
ggcgacacaa agcacatgat tttcttacia cggagataaa accaaaaaaa tatttcatga	7920
acaacctaga acaataaag cttttatata ataaatatat aaataaataa aggctatgga	7980
ataatatact tcaatatatt tggattaaat aaattgttgg cggggttgat atatttatac	8040
acacctaaag tcacttcaat ctcaatttca cttaactttt attttttttt tctttttatt	8100
tatcataaag agaattttga taataacttt tttaacatat ttttatgaca ttttttattg	8160
gtgaaaactt attaaaaac ataaattttg taagttagat ttattttaaag agttcctctt	8220
cttattttta attttttaat aaatttttaa ataactaaaa tttgtgttaa aaatgttaa	8280
aaagtgtgtt attaaccttt ctcttcgagg atccgtacga tcccacatgc aagtttttat	8340

-continued

---

```

ttcaatccct tttccttga ataactgacc aagaacaaca agaaaaaaaa aaaaaaagaa 8400
aaggatcatt ttgaaaggat atttttcgct cctattcaaa tactgtatth ttaccaaaaa 8460
aactgtatth ttcctacact ctcaagctth gtttttcgct tcgactctca tgatttcctt 8520
catatgccaa tcactctatt tataaatggc ataaggtagt gtgaacaatt gcaaagcttg 8580
tcatcaaaag cttgcaatgt acaaatatg gtttttcctg cctttcaaaa ttatctgcac 8640
cccctagcta ttaatctaac atctaagtaa ggctagtgaa ttttttcgaa tagtcatgca 8700
gtgcattaat ttccccgta ctatthtggc tttgactcca aactggccc cgtacatccg 8760
tcccctatta catgaaaaga aatattgth atattctaa ttaaaaatat tgtcccttct 8820
aaatthtcat atagthtaatt attatattac ttttttctct attctattag ttctatthtc 8880
aaattattat ttatgcatat gtaagtaga ttataththt gctatatact taaatathtc 8940
taaattatta aaaaaagact gatatgaaa atthattctt tttaaagcta tatcatttha 9000
tatatactth tctthtctct tctthtcaat ttctattcaa ttaataaga aataaththt 9060
gtaaaththt atthtcaat ttataaaaat atthtactth atatgththt tcacaththt 9120
gttaaaaaa tcatatcatt atgattgaa gagaggaaat tgacagttag taataagtag 9180
tgagaaaaa atgtgttatt tcttaaaaaa aacctaaaca aacatgtatc tactctctat 9240
ttcatctatc tctcatttca tttttctctt tatctcttcc tttathththt tatcatatca 9300
tttcacatta atthththta ctctcttht tttttctctc tatccctctc thatttccac 9360
tcatatatac actccaaaat tggggcatgc ctttatcact actctatctc ctccactaaa 9420
tcattthaaat gaaactgaaa agcattggca agtctctctc cctctcaag tgatttccaa 9480
ctcagcattg gcctctaatt gattcagtag atctattgca tgtgtaaaag tctttccaca 9540
atacataact atthaaatct cttaaataaa taaaggataa aatathththt thtctcata 9600
aaathaaat atgtaththt thgtthtagat gtatattcga ataaatctaa atathgata 9660
atgathththt atathgata aacatataat caatathaa tatgaththt ththtatag 9720
gttgtagaca thaththata aggataaaaa atathgataa aataaththt aathaththt 9780
atathttaga gaaaaaaaa thththtagcc ataaataaat gaccagcata ththtacaacc 9840
thtagthattc ataaatctct atathgata thgaaathaa aaacagataa tcgthaaagg 9900
aaggathctc acgtctctc thgctattg ththtctatgc aaacagaaag ggacgaaaa 9960
ccactctacc atgathcact cttcacacca ththtactag caaacaagtc tcaacaactg 10020
aagccagctc tctthccgth tctththtaca acactthctt gaaatagta gaththththt 10080
thcactgath thathaaact gccaaaagat gctthattgaa tagagtgcac atthgthaatg 10140
tactactaat tagaacatga aaaagcattg thctaacacg ataatctgt gaaggcgtta 10200
actccaaaga tccaathtca ctataaat thtgacgaaa gcaaatgaa thcactagc 10260
tgagagagaa aggaaaggth aactaagaag caactctca gcggccgcat gagccgtaaa 10320
ggthcaatc aacagtgct thgtthctta gggacaagca thgtacttht gthgathct 10380
gthgaaacct gagthtctca cgtthtacta atgthaaagg caaaaaataa acacagaa 10440
agthcgttht thtcaataa thgthaaagg gaaatggaa ccatgcctcc thtctthgath 10500
gthgaththaa atathgacg atgthacc 10528

```

&lt;210&gt; SEQ ID NO 31

&lt;211&gt; LENGTH: 11721

&lt;212&gt; TYPE: DNA

-continued

---

```

<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: vector pKR1147
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: (7362)..(7362)
<223> OTHER INFORMATION: n is a, c, g, or t

<400> SEQUENCE: 31
ggccgcatga gccgtaaagg ttcaatacaa cgagtgcttg ttttcttagg gacaagcatt    60
gtacttatgt atgattctgt gtaaccatga gtcctccacg ttgtactaat gtgaagggca    120
aaaataaaac acagaacaag ttcggttttc tcaaataatg tgaaggtaga aaatggaacc    180
atgcctcctc tcttgcatgt gatttaaaat attagcagat ggtaccgtac gagatccggc    240
cggccagatc ctgcaggaga tccaagcttg gcgcgccgtt ctatagtgtc acctaaatcg    300
tatgtgatg atacataaagg ttatgtatta attgtagccg cgttctaacy acaatatgtc    360
catatgggtg actctcagta caatctgtc tgatgccgca tagttaagcc agccccgaca    420
cccgccaaca cccgctgacg cgccctgacg ggcttgtctg ctcccggcat ccgcttacag    480
acaagctgtg accgtctccg ggagctgcat gtgtcagagg ttttcaccgt catcaccgaa    540
acgcgcgaga cgaaggggcc tcgtgatacg cctattttta taggtaatg tcatgaccaa    600
aatcccttaa cgtgagtttt cgttccactg agcgtcagac cccgtagaaa agatcaaagg    660
atcttcttga gatccttttt ttctgcgctt aatctgctgc ttgcaaaaa aaaaaccacc    720
gtaccagcgg gtggtttggt tgccggatca agagetacca actcttttcc cgaaggtaac    780
tggcttcagc agagcgcaga taccaaatac tgtccttcta gtgtagccgt agttaggcca    840
ccacttcaag aactctgtag caccgcctac atacctcgtc ctgctaatac tgttaccagt    900
ggctgctgcc agtggcgata agtctgtctc taccgggttg gactcaagac gatagttacc    960
ggataaggcg cagcggctgg gctgaacggg gggttcgtgc acacagccca gcttggagcg   1020
aacgacctac accgaactga gatacctaca gcgtgagcat tgagaaagcg ccacgettcc   1080
cgaaggggaga aaggcggaca ggtatccggg aagcggcagg gtcggaacag gagagcgcac   1140
gagggagctt ccagggggaa acgcctggta tctttatagt cctgtcgggt ttccccacct   1200
ctgacttgag cgtcgatttt tgtgatgctc gtcagggggg cggagcctat ggaaaaacgc   1260
cagcaacgcg gcctttttac ggttccctgg cttttgctgg ccttttgctc acatgttctt   1320
tcctgcgtta tcccctgatt ctgtggataa ccgattacc gcctttgagt gagctgatac   1380
cgctcgccgc agccgaacga ccgagcgcag cgagtcagtg agcgaggaag cggaaagagc   1440
cccaatacgc aaaccgcctc tcccgcgcg ttggccgatt cattaatgca ggttgatcga   1500
ttcgacatcg atctagtaac atagatgaca ccgcgcgcga taatttatcc tagtttgccg   1560
gctatatttt gttttctatc gcgtattaaa tgtataattg cgggactcta atcataaaaa   1620
cccatctcat aaataacgtc atgcattaca tgtaattat tacatgctta acgtaattca   1680
acagaaatta tatgataatc atcgaagac cggcaacagg attcaatctt aagaaacttt   1740
attgccaaat gtttgaacga tctgcttcca cgcactcctt ctttaggtac ctactattc   1800
ctttgccctc ggacgagtgc tggggcgtcg gtttccacta tcggcgagta cttctacaca   1860
gccatcggtc cagacggccg cgcttctgcg ggcgatttgt gtaacgccga cagtcccggc   1920
tccggatcgg acgattcgtc cgcategacc ctgcgcccaa gctgcatcat cgaattgccc   1980
gtcaaccaag ctctgataga gttggccaag accaatgcgg agcatatacg cccggagccg   2040
cggcgatcct gcaagctccg gatgcctccg ctcgaagtag cgcgtctgct gctccataca   2100

```

-continued

---

agccaaccac	ggcctccaga	agaagatggt	ggcgacctcg	tattgggaat	ccccgaacat	2160
cgctctgctc	cagtcaatga	ccgctgttat	gcggccattg	tccgtcagga	cattggttga	2220
gccgaaatcc	gcgtgcacga	ggtgccggac	ttcggggcag	tcctcgccc	aaagcatcag	2280
ctcatcgaga	gcctgcgoga	cggacgcact	gacggtgtcg	tccatcacag	tttgccagtg	2340
atacacatgg	ggatcagcaa	tcgcgcataat	gaaatcacgc	catgtagtgt	attgaccgat	2400
tccttgccgt	ccgaatgggc	cgaaccogct	cgtctggcta	agatcgccc	cagcgatcgc	2460
atccatggcc	tccgcgaccg	gctgcagAAC	agcgggcagt	tcggtttcag	gcaggctttg	2520
caacgtgaca	ccctgtgcac	ggcgggagat	gcaataggtc	aggctctcgc	tgaattcccc	2580
aatgtcaagc	acttccgga	tcgggagcgc	ggccgatgca	aagtgccgat	aaacataacg	2640
atctttgtag	aaaccatcgg	cgcagctatt	taccgcagg	acatatccac	gccctctac	2700
atcgaagctg	aaagcacgag	attcttcgcc	ctccgagagc	tgcacaggt	cggagacgct	2760
gtcgaacttt	tcgatcagaa	acttctcgac	agacgtcgcg	gtgagttcag	gctttttcat	2820
ggtttaataa	gaagagaaaa	gagttctttt	gttatggctg	aagtaataga	gaaatgagct	2880
cgagcgtgtc	ctctccaaat	gaaatgaact	tccttatata	gaggaaggt	cttgccaagg	2940
atagtgggat	tgtcgcgcat	cccttacgtc	agtggagatg	tcacatcaat	ccacttgctt	3000
tgaagacgtg	gttggaacgt	cttcttttct	cacgatgctc	ctcgtgggtg	gggggccatc	3060
tttgggacca	ctgtcggcag	aggcatcttg	aatgatagcc	tttctttat	cgcaatgatg	3120
gcattttag	gagccacctt	ccttttctac	tgtcctttcg	atgaagtgac	agatagctgg	3180
gcaatggaat	ccgaggaggt	ttcccgaat	tatcctttgt	tgaaaagtct	caatagccct	3240
ttggtctctc	gagactgtat	ctttgacatt	tttgagtag	accagagtgt	cgtgctccac	3300
catgttgacg	aagattttct	tcttgtcatt	gagtcgtaa	agactctgta	tgaactgttc	3360
gccagtcttc	acggcgaggt	ctgttagatc	ctcgatttga	atcttagact	ccatgcattg	3420
ccttagattc	agtaggaact	acctttttag	agactccaat	ctctattact	tgccttggtt	3480
tatgaagcaa	gccttgaatc	gtccatactg	gaatagtact	tctgatcttg	agaaatagt	3540
ctttctctgt	gttcttgatg	caattagtcc	tgaatctttt	gactgcatct	ttaaccttct	3600
tgggaaggta	tttgatctcc	tggagattgt	tactcgggta	gatcgtcttg	atgagacctg	3660
ctgcgtaggc	ctctctaacc	atctgtgggt	cagcattctt	tctgaaattg	aagaggctaa	3720
ccttctcatt	atcagtgggt	aacatagtgt	cgtaaccttc	accttcgaac	ttccttctca	3780
gatcgtaaag	atagaggaaa	tcgtccattg	taatctccgg	ggcaaaggag	atctcttttg	3840
gggctggatc	actgctgggc	cttttggttc	ctagcgtgag	ccagtgggct	ttttgctttg	3900
gtgggcttgt	tagggcctta	gcaaagctct	tgggcttgag	ttgagcttct	cctttgggga	3960
tgaagtcaa	cctgtctggt	tgtgacttg	ttgtgtacgc	gtcagctgct	gctcttgccct	4020
ctgtaatagt	ggcaaatctt	ttgtgtgcaa	ctccgggaac	gccgtttggt	gccgcctttg	4080
tacaacccca	gtcatcgat	ataccggcat	gtggaccgtt	atacacaacg	tagtagttga	4140
tatgaggggtg	ttgaataacc	gattctgctc	tgagaggagc	aactgtgctg	ttaaagctcag	4200
atctttgtgg	gattggaatt	ggatcgatct	cgatcccgcg	aaattaatac	gactcactat	4260
agggagacca	caacggtttc	cctctagaaa	taattttggt	taactttaag	aaggagatat	4320
acctatggaa	aagcctgaac	tcaccgcgac	gtctgtcgag	aagtttctga	tcgaaaagtt	4380
cgacagcgtc	tccgacctga	tgcagctctc	ggagggcgaa	gaatctcgtg	ctttcagctt	4440

-continued

---

cgatgtagga	gggcgtggat	atgtcctgcg	ggtaaaatagc	tgcgccgatg	gtttctacaa	4500
agatcgttat	gtttatcggc	actttgcatc	ggccgcgctc	ccgattccgg	aagtgcttga	4560
cattggggaa	ttcagcgaga	gocctgacct	ttgcatctcc	cgccgtgcac	agggtgtcac	4620
gttgcaagac	ctgcctgaaa	ccgaactgcc	cgctgttctg	cagccggctc	cggaggctat	4680
ggatgcgac	gctgcggccg	atcttagcca	gacgagcggg	ttcggcccat	tcggaccgca	4740
aggaatcggg	caatacacta	catggcgtga	tttcatatgc	gcgattgctg	atccccatgt	4800
gtatcactgg	caaaactgtg	tggacgacac	cgctcagtgc	tccgtcgcgc	aggctctcga	4860
tgagctgatg	ctttgggccc	aggactgcc	cgaagtccgg	cacctcgtgc	acgcgattt	4920
cggctccaac	aatgtcctga	cggacaatgg	ccgcataaca	gcggtcattg	actggagcga	4980
ggcgatgttc	ggggattccc	aatacagagt	cgccaacatc	ttcttctgga	ggccgtgggt	5040
ggcttgatg	gagcagcaga	cgcgctactt	cgagcggagg	catccggagc	ttgcaggatc	5100
gcccgcgctc	cgggcgtata	tgtcccgcat	tggctttgac	caactctatc	agagcttggg	5160
tgacggcaat	ttcgatgatg	cagcttgggc	gcagggtcga	tgcgacgcaa	tcgtccgatc	5220
cggagccggg	actgtcgggc	gtacacaaat	cgcccgcaga	agcgcggccg	tctggaccga	5280
tggctgtgta	gaagtactcg	ccgatagtgg	aaaccgacgc	cccagcactc	gtccgagggc	5340
aaaggaatag	tgaggtacag	cttgatcga	tccggctgct	aacaaagccc	gaaaggaagc	5400
tgagttggct	gctgccaccg	ctgagcaata	actagcataa	ccccttgggg	cctctaaacg	5460
ggcttgagg	ggttttttgc	tgaaggagg	aactatatcc	ggatgatcgg	gcgcccgctc	5520
gacggatcca	ctagttctag	agcggcccgc	gccgtcgacg	gatataatga	gccgtaaaca	5580
aagatgatta	agtagtaatt	aatcgtact	agtaaaagtg	gcaaaagata	acgagaaaga	5640
accaattttc	ttgcattcgg	ccttagcggg	aggcatatat	aagctttgat	tattttattt	5700
agtgtaatga	tttcgtacaa	ccaaagcatt	tatttagtac	tctcactctt	gtgtcggcgc	5760
cggccgctac	aggaacaggt	ggtggcggcc	ctcggcgcgc	tcgtactgct	ccacgatggg	5820
gtagtcctcg	ttgtgggagg	tgatgtccag	cttgaggctc	acgtagtagt	agccgggag	5880
ctgcacgggc	ttcttgcca	tgtagatgga	cttgaactcc	accaggtagt	ggccgcccgc	5940
cttcagcttc	aggcccttgt	ggatctcgcc	cttcagcagc	ccgtcgcggg	ggtacaggcg	6000
ctcggtgagg	gcctcccagc	ccatagtctt	cttctgcatt	acggggccgt	cggaggggaa	6060
gttcacgccc	atgaacttca	cctttagatg	gaaggagccc	tcctgcaggg	aggagtctcg	6120
ggtcacggtc	accacgcgcg	cgtcctcgaa	gttcatcacg	cgctcccact	tgaagccctc	6180
ggggaaggac	agcttcttgt	agtcggggat	gtcggcgggg	tgttccacgt	acaccttggg	6240
gccgtactgg	aactgggggg	acaggatgtc	ccaggcgaag	ggcagggggc	cgcccttggg	6300
caccttcagc	ttggcgtctc	gggtgccctc	gtagggggcg	ccctcgcctc	cgccctcgat	6360
ctcgaactcg	tggccgttca	cggagccctc	catgcccacc	ttgaagcga	tgaactcctt	6420
gatgacgtcc	tcggaggagg	ccatgggccc	cttggggggc	tatggaagac	tttcttagtt	6480
agttgtgtga	ataagcaatg	ttgggagaat	cgggactact	tataggatag	gaataaaaca	6540
gaaaagtatt	aagtgctaat	gaaatattta	gactgataat	taaaatcttc	acgtatgtcc	6600
acttgatata	aaaacgtcag	gaataaagga	agtacagtag	aatttaaagg	tactcttttt	6660
atatataccc	gtgttctctt	tttggctagc	tagttgcata	aaaaataatc	tatatatttt	6720
tcattatttt	aaatatctta	tgagatggta	aatattttatc	ataatttttt	ttactattat	6780
ttattatttg	tgtgtgtaat	acatatagaa	gttaattaca	aattttattt	actttttcat	6840

-continued

---

tattttgata	tgattcaoca	ttaatttagt	gttattattt	ataatagttc	attttaactc	6900
ttttgtatat	attatgcgtg	cagtactttt	ttcctacata	taactactat	tacattttat	6960
ttatataata	tttttattaa	tgaattttcg	tgataaatg	taatattggt	cattattatt	7020
tcagattttt	taaaaatatt	tgtgttatta	tttatgaaat	atgtaatttt	tttagtattt	7080
gattttatga	tgataaagtg	ttctaaattc	aaaagaaggg	ggaaagcgta	aacattaaaa	7140
aacgctcatca	aacaaaaaca	aaatcttggt	aataaagata	aaactgtttg	ttttgatcac	7200
tgttattttcg	taataataaa	acattattta	tatttatatt	gttgacaacc	aaatttgcct	7260
atcaaatcta	accaatataa	tgcatgogtg	gcaggtaatg	tactaccatg	aacttaagtc	7320
atgacataat	aaaccgtgaa	tctgaccaat	gcatgtacct	anctaaattg	tattttgtgac	7380
acgaagcaaaa	tgattcaatt	cacaatggag	atgggaaaaa	aataatgaag	aaccagaaac	7440
taagaaagct	tttctgaaaa	ataaaataaa	ggcaatgtca	aaagtatact	gcatcatcag	7500
tccagaaagc	acatgatatt	tttttatcag	tatcaatgca	gctagtttta	ttttacaata	7560
tcgatatagc	tagtttaaat	atattgcagc	tagattttata	aatattttgtg	ttattattta	7620
tcattttgtg	aatcctgttt	ttagtatttt	agtttatata	tgatgataat	gtattccaaa	7680
tttaaagaa	gggaaataaa	tttaaacaag	aaaaaaaaagtc	atcaaaaaaa	aaacaaatga	7740
aagggtgga	agatgttacc	atgtaatgtg	aatgttacag	tattttctttt	attatagagt	7800
taacaaatta	actaatatga	ttttgttaat	aatgataaaa	tattttttttt	attattattt	7860
cataatataa	aaatagttta	cttaatataa	aaaaaaaaattc	atcgttcaca	acaaagttag	7920
ccacctaat	taacctgca	tgtacctcatg	gaccatatta	ggtaaccatc	aaacctgatg	7980
aagagataaa	gagatgaaga	cttaagtcac	aacacaaaaac	cataaaaaaac	aaaaatataa	8040
tcaaccgtca	atctgaccaa	tgcatgaaaa	agctgcaata	gtgagtggcg	acacaaagca	8100
catgattttc	ttacaacgga	gataaaacca	aaaaaatatt	tcatgaacaa	cctagaacaa	8160
ataaagcttt	tatataataa	atatataaat	aaataaaggc	tatggaataa	tatacttcaa	8220
tatatttgga	ttaaataaat	tgttggcggg	gttgatatat	ttatacacac	ctaaagtcc	8280
ttcaatctca	ttttcactta	actttttatt	tttttttctt	tttattttatc	ataaagagaa	8340
tattgataat	atacttttta	acatattttt	atgacatttt	ttattggtga	aaacttatta	8400
aaaatcataa	atttttgaag	ttagattttat	ttaaagagtt	cctctttctta	ttttaaattt	8460
tttaataaat	ttttaaataa	ctaaaatttg	tgttaaaaat	gttaaaaaag	tgtgttatta	8520
acccttctct	tcgaggatcc	gtacgatccc	acatgcaagt	ttttatttca	atcccttttc	8580
ctttgaataa	ctgaccaaga	acaacaagaa	aaaaaaaaaaa	aaagaaaagg	atcattttga	8640
aaggatattt	ttcgcctcta	ttcaaatact	gtattttttac	caaaaaaact	gtatttttcc	8700
tacactctca	agctttgttt	ttcgcctcga	ctctcatgat	ttccttcata	tgccaatcac	8760
tctatttata	aatggcataa	ggtagtgtga	acaattgcaa	agcttgtcat	caaaagcttg	8820
caatgtacaa	attaatgttt	ttcatgcctt	tcaaaattat	ctgcaccccc	tagctattaa	8880
tctaacatct	aagtaaggct	agtgaatttt	ttcgaatagt	catgcagtgc	attaatttcc	8940
ccgtgactat	tttggctttg	actccaacac	tggtcccgta	catccgtccc	tcattacatg	9000
aaaagaaata	ttgtttatat	tcttaattaa	aaatattgtc	ccttctaaat	tttcatatag	9060
ttaattatta	tattactttt	ttctctatcc	tattagttct	attttcaaat	tattattttat	9120
gcatatgtaa	agtacattat	atttttgcta	tatacttaaa	tattttctaaa	ttattaaaaa	9180

-continued

---

aagactgata	tgaaaaat	ttcttttt	aagctat	atattat	tacttttt	9240
tttctttt	ttcatttt	attcaatt	ataagaa	aattttgt	atttttat	9300
atcaattt	aaaaat	tactttat	gttttttc	atttttgt	aacaaatc	9360
atcattat	tgaaag	ggaattg	agtgag	aagtgat	aaaaaat	9420
gttatttc	aaaaaa	taaacaaa	tgtatct	ctctatt	tctatct	9480
atctcatt	tctcttt	tctttctt	tttttttc	atctcatt	acattaat	9540
tttttact	ctttatt	tctctct	cctctct	ttccact	atatact	9600
caaaattg	gcattgc	atcactac	tatctct	actaaat	ttaaatg	9660
ctgaaa	gtgcaag	tctctct	ctcaagt	ttccaac	gcattgg	9720
ctaattg	cagtata	attgcatt	taaaagt	tccaca	ataact	9780
attaact	tataat	ggataaa	tttttttc	ttcata	taaaat	9840
tatttttt	ttagatt	attcgaa	atctaat	atgata	ttttttat	9900
tgattaaa	tataat	atataat	atatttt	atatagg	tacacata	9960
tttata	gataaa	aatttt	atatttt	ttacgag	10020	
aaaaaat	ttagcc	ataaat	agcatatt	acaacct	taattca	10080
attcct	gtattt	aattaa	agataat	taagga	aatcct	10140
catctct	catttgt	tcatgca	agaaagg	gaaaaac	ctcacc	10200
atcact	acaccat	tactagc	caagtct	caactga	cagctct	10260
tccgttt	tttaca	tttcttg	atagtag	ttttttt	catgatt	10320
taacgtg	aaagatt	attgaat	gtgcac	gtaatgt	actaatt	10380
acatgaaa	gcattgt	aacacga	tctgtga	gcgttaa	caaagat	10440
atctcact	ataaat	acgaaag	aatgaat	catagct	agagaa	10500
aaggtt	aagaag	acttcag	cgcattg	agatct	ggcagt	10560
tccgcc	tcgcc	cctcct	ctcctct	gcgacac	tctcgt	10620
tcccgaa	aggcgg	cgcgac	caaggcc	gccgac	ataag	10680
cgcgaag	cccgcg	cgcgcg	gaagagg	tccgtct	ggggag	10740
caggcac	tgacgg	ggttcg	gcattct	gacaagc	gcctcgc	10800
gctccac	aagaaga	gcaggca	ctacct	gcgtat	gcgagg	10860
agctgct	gcctat	tcgcag	caagtac	ggctct	ctctgct	10920
cttccct	gtgatt	ccagcag	gccggag	gaggcct	cccggg	10980
gtacctg	tccctc	gcaggag	cggctct	aggggc	ccaagt	11040
aggcgt	aggcat	acaacgg	gtgggag	cggatt	gagtc	11100
gaacaag	ctctact	gaacatt	cactca	gaggcag	aggcct	11160
ccttgcc	attgaat	gtggcgt	tgctgta	aactcg	tcagct	11220
cctggac	ccgctgt	tggcac	ccaacag	ccacagg	tgcggc	11280
caacca	cctcaac	atcagag	aaccgga	acagag	agccgg	11340
aagcga	aagacac	atggcag	agaacct	gagaac	gcctgac	11400
caccg	cccctca	cagtcag	cagcat	gaggcct	ggagcc	11460
catggat	gagctag	ccatgt	accaaact	ggcagct	tcaact	11520
cgagtgt	gctgacg	acttcag	caacat	tgctgt	atgggt	11580

-continued

---

```

tcggcgtgac gaaggaagca aggatggtgt aggtctggca gatttcagtc tgtttgaggc 11640
aggtgatgtc cagctgaagg atgttcttcc ggatatggaa gaggggatac aacctccagc 11700
gatgatcagt gttgcaacg c 11721

```

```

<210> SEQ ID NO 32
<211> LENGTH: 19713
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: vector pKR1220
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: (15083)..(15083)
<223> OTHER INFORMATION: n is a, c, g, or t

```

```

<400> SEQUENCE: 32

```

```

cgcgccagat cctctagagt cgacctgcag gcatgcaagc ttggcgtaat catggtcata 60
gctgtttcct gtgtgaaatt gttatccgct cacaattcca cacaacatac gagccggaag 120
cataaagtgt aaagcctggg gtgcctaatt agtgagctaa ctcacattaa ttgctgtgag 180
ctcactgccc gctttccagt cgggaaacct gtcgtgccag ctgcattaat gaatcggcca 240
acgcgcgggg agaggcgggt tgcgtattgg atcgatccct gaaagcgacg ttggatgta 300
acatctacaa attgcctttt cttatcgacc atgtacgtaa gcgcttacgt ttttgggtga 360
cccttgagga aactggtagc tgttgtgggc ctgtggtctc aagatggatc attaatttcc 420
accttcacct acgatggggg gcatcgcaac ggtgagtaat attgtacggc taagagcgaa 480
tttggcctgt agacctcaat tgcgagcttt ctaatttcaa actattcggg cctaactttt 540
ggtgtgatga tgctgactgg caggatatac accgttgtaa tttgagctcg tgtgaataag 600
tcgctgtgta tgtttgtttg attgttctg ttggagtgca gccatttca cgggacaagt 660
cggetagatt gatttagccc tgatgaactg ccgaggggaa gccatcttga gcgcggaatg 720
ggaatggatt tcgttgtaca acgagacgac agaaccacca cgggaccgag cttcgcgagc 780
ttttgtatcc gtggcatcct tggccgggc gatttgttca cgtccatgag gcgctctcca 840
aaggaaacgca tattttccgg tgaaccttt cgggttcttc ctctactcga cctcttgaag 900
tcccagcatg aatgttcgac cgctccgcaa gcgatcttt ggcgcaacca gccggtttcg 960
cacgtcgatt ctcgcgagcc tgcatacttt ggcaagattg ctgaatgacg ctgatgctt 1020
atcgcaatct gcgataatgg ggtaaagtac cgggtgaaggc cgcaggtcag gccgcctgag 1080
cactcagtgt cttggatgtc cagttccacg gcagctgttg ctcaagcctg ctgatcggag 1140
cgtccgcaag gtcggcgcgg acgtcggcaa gccaggcctg cggatcgatg ttattgagct 1200
tggcgctcat gatcagtgtc gccatgaacg ccgcacgttc agcacaacga tccgatccgg 1260
caaacagcca tgacttccctg ccgagtacat agcctctgag cgttcgttcg gcagcattgt 1320
tcgtcaggca aatcgggccc tcatcgagga atgacgtaat gccatcccat cgcttgagca 1380
tgtaatttat cgctcggcg acgggagaaac tgcgcgacaa tttccccgc tgggtttcga 1440
gccaatcatg cagctcttcg gcgagtgacc ttgatcaggc caccgccacg acccggaag 1500
acgaacagat gcctgcgcat cggatcgcgc ttcagcgtct cttgcacct cagcgcacaa 1560
ccgggaaagc ctttgcgcat gtccgtactt atgtcgccac ttgggagggc ttcgtctacg 1620
tggccttcgt gatcagctc ttcgccctc gcattgtcgg atggcggcg agccggacag 1680
cacatgcagg ctttgcctc gatgccctc aggaggctca tcatgatcgg cgtcccctc 1740

```

-continued

---

atggcggcct	agtgcacac	tcggatcgcg	gtgttcaata	cgtgtccttt	cgctattccg	1800
agcggttggc	agaagcaggt	atcgagccat	ctatcggaag	cgteggcgac	agcacgacaa	1860
cgccctcgca	gaagcgatca	acggctctta	caaggccgag	gtcattcacc	ggcgtggacc	1920
atggaggagc	ttcgaagcgg	tcgagttcgc	taccttgaa	tggatagact	ggttcaacca	1980
cggcggcttt	tgaagcccat	cggcaatata	cggccagccg	aagacgagga	tcagtattac	2040
gccatgctgg	acgaagcagc	catggctgcg	cattttaacg	aatggcctc	cggcaaaccc	2100
ggtgcggttc	acttgttgcg	tgggaaagt	cacgggactc	cgcgcacgag	ccttctctgt	2160
aatagccata	tcgaccgaat	tgacctgcag	gggggggggg	gaaagccacg	ttgtgtctca	2220
aaatctctga	tgttacattg	cacaagataa	aaatatatca	tcatgaacaa	taaaactgtc	2280
tgcttacata	aacagtaata	caaggggtgt	tatgagccat	attcaacggg	aaacgtcttg	2340
ctcggaggcg	cgattaaatt	ccaacatgga	tgctgattta	tatgggtata	aatgggctcg	2400
cgataaatgtc	gggcaatcag	gtgcgacaat	ctatcgattg	tatgggaage	ccgatgcgcc	2460
agagttgttt	ctgaaacatg	gcaaaggtag	cgttgccaat	gatgttacag	atgagatggt	2520
cagactaac	tggtgcagcg	aatttatgcc	tcttccgacc	atcaagcatt	ttatccgtac	2580
tctctgatgat	gcattggttac	tcaccactgc	gatccccggg	aaaacagcat	tccaggtatt	2640
agaagaatat	cctgatccag	gtgaaaatat	tgttgatgcg	ctggcagtgt	tctgcgccg	2700
gttgctctcg	attcctgttt	gtaattgtcc	ttttaacagc	gatcgcgtat	ttcgtctcgc	2760
tcaggcgcaa	tcacgaatga	ataacggttt	ggttgatgcg	agtgattttg	atgacgagcg	2820
taatggctgg	cctgttgaac	aagtctggaa	agaaatgcat	aagcttttgc	cattctcacc	2880
ggattcagtc	gtcactcatg	gtgatttttc	acttgataac	cttatttttg	acgaggggaa	2940
attaataggt	tgtattgatg	ttggacgagt	cggaatcgca	gaccgatacc	aggatcttgc	3000
catcctatgg	aactgcctcg	gtgagttttc	tccttcatta	cagaaacggc	tttttcaaaa	3060
atattggtatt	gataatcctg	atatgaataa	attgcagttt	catttgatgc	tcgatgagtt	3120
tttctaatca	gaattgggta	attgggtgta	acactggcag	agcattacgc	tgacttgacg	3180
ggacggcggc	tttgttgaat	aaatcgaact	tttgcctgag	tgaaggatca	gatcacgcat	3240
cttcccagca	acgcagaccg	ttccgtggca	aagcaaaagt	tcaaaatcac	caactgggtcc	3300
acctacaaca	aagctctcat	caaccgtggc	tcctcactt	tctggctgga	tgatggggcg	3360
attcaggcct	ggtatgagtc	agcaaacact	tcttcacgag	gcagacctca	gcgccccccc	3420
ccccctcgag	gtcttttcca	atgatgagca	cttttaaagt	tctgctatgt	ggcgcgggtat	3480
tatcccgtgt	tgacgccggg	caagagcaac	tcggctgcgg	catacactat	tctcagaatg	3540
acttgggtga	gtactcacca	gtcacagaaa	agcatcttac	ggatggcatg	acagtaagag	3600
aattatgcag	tgctgccata	accatgagtg	ataaacactgc	ggccaactta	cttctgacaa	3660
cgatcggagg	accgaaggag	ctaaccgctt	ttttgcacaa	catgggggat	catgtaactc	3720
gccttgatcg	ttgggaaccg	gagctgaatg	aagccatacc	aaacgacgag	cgtgacacca	3780
cgatgcctgt	agcaatggca	acaacgttgc	gcaaacattt	aactggcgaa	ctacttactc	3840
tagcttcccg	gcaacaatta	atagactgga	tggaggcggga	taaagttgca	ggaccacttc	3900
tgcgctcggc	ccttccggct	ggctggttta	ttgctgataa	atctggagcc	ggtgagcgtg	3960
ggtctcggcg	tatcattgca	gcactggggc	cagatggtaa	gccctcccg	atcgtagtta	4020
tctacacgac	ggggagtcag	gcaactatgg	atgaacgaaa	tagacagatc	gctgagatag	4080
gtgcctcact	gattaagcat	tgtaactgt	cagaccaagt	ttactcatat	atactttaga	4140

-continued

---

ttgatttaaa	acttcatttt	taatttaaaa	ggatctaggt	gaagatcctt	tttgataatc	4200
tcatgaccaa	aatcccttaa	cgtgagtttt	cgttccactg	agcgtcagac	cccgtagaaa	4260
agatcaaagg	atcttcttga	gatccctttt	ttctgocgct	aatctgctgc	ttgcaaaaa	4320
aaaaaccacc	gctaccagcg	gtggtttggt	tgccggatca	agagctacca	actctttttc	4380
cgaaggtaac	tggcttcagc	agagcgcaga	taccaaatat	tgctcttcta	gtgtagccgt	4440
agttaggcca	ccacttcaag	aactctgtag	caccgcctac	atacctcgct	ctgctaatac	4500
tgttaccagt	ggctgctgcc	agtggcgata	agtcgtgtct	taccgggttg	gactcaagac	4560
gatagttaac	ggataaggcg	cagcggtcgg	gctgaacggg	gggttcgtgc	acacagccca	4620
gcttgagcgc	aacgacctac	accgaactga	gatacctaca	gcgtgagcta	tgagaaagcg	4680
ccacgcctcc	cgaagggaga	aagggcgaca	ggtatccggt	aagcggcagg	gtcggaacag	4740
gagagcgcac	gagggagcct	ccaggggaa	acgcctggta	tctttatagt	cctgtcgggt	4800
ttcgccacct	ctgacttgag	cgctcatttt	tgtgatgctc	gtcagggggg	cggagcctat	4860
ggaaaaacgc	cagcaacgcg	gcctttttac	ggttcctggc	cttttgetgg	ccttttgctc	4920
acatgttctt	tcctgcgtta	tcccttgatt	ctgtggataa	ccgtattacc	gcctttgagt	4980
gagctgatac	cgctcgcgcg	agccgaacga	ccgagcgcag	cgagtcagtg	agcgaaggag	5040
cggaaagagcg	cctgatgcgg	tattttctcc	ttacgcctct	gtgcgggtatt	tcacaccgca	5100
tatggtgcac	tctcagtaca	atctgctctg	atgccgcata	gttaagccag	tatacactcc	5160
gctatcgcta	cgtgactggg	tcattgctgc	gccccgacac	ccgccaacac	ccgctgacgc	5220
gccctgacgg	gcttgtctgc	tcccggcctc	cgcttacaga	caagctgtga	ccgtctccgg	5280
gagctgcatg	tgtcagaggt	ttcacccgtc	atcacgaaaa	cgcgcgaggc	agggtgccct	5340
gatgtgggcg	ccggcggctg	agtggcgacg	gcgcggcctg	tcgcgcacct	ggtagattgc	5400
ctggccgtag	gccagccatt	tttgagcggc	cagcggcgcg	gataggccga	cgcgaagcgg	5460
cggggcgtag	ggagcgcagc	gaccgaaggg	taggcgcttt	ttgcagctct	tcggctgtgc	5520
gctggccaga	cagttatgca	caggccaggc	gggttttaag	agttttaata	agttttaaag	5580
agttttaggg	ggaaaaatcg	ccttttttct	cttttatatc	agtcacttac	atgtgtgacc	5640
ggttcccaat	gtacggcttt	gggttcccaa	tgtacgggtt	ccggttccca	atgtacggct	5700
ttgggttccc	aatgtacgtg	ctatccacag	gaaagagacc	ttttcgacct	ttttcccctg	5760
ctagggcaat	ttgcctagc	atctgctccg	tacattagga	accggcggat	gcttgcacct	5820
cgatcagggt	gcggtagcgc	atgactagga	tcgggccagc	ctgccccgcc	tcctccttca	5880
aatcgtactc	cggcaggtca	tttgacccga	tcagcttgcg	cacggtgaaa	cagaacttct	5940
tgaactctcc	ggcgcgtcca	ctcgttctgt	agatcgtctt	gaacaacctt	ctggcttctg	6000
ccttgctctc	ggcgcggcgt	gccagcgggt	agagaaaaac	gccgatgccg	ggatcgatca	6060
aaaagtaatc	ggggtgaaac	gtcagcacgt	ccgggttctt	gccttctgtg	atctcgcggg	6120
acatccaatc	agctagctcg	atctcgatgt	actccggccg	ccgggtttcg	ctctttacga	6180
tctttagagc	gctaatacaag	gcttcaacct	cggataacct	caccaggcgg	ccgttcttgg	6240
ccttctctct	acgctgcatg	gcaacgtcgc	tggtgtttaa	ccgaatgcag	gtttctacca	6300
ggctcgtctt	ctgctttccg	ccatcggctc	gccggcagaa	cttgagtacg	tcgcgaacgt	6360
gtggacggaa	cacgcggcgc	ggcttctctc	ccttcccttc	ccggtatcgg	tcatggatt	6420
cggttagatg	ggaaaaccgc	atcagtaacca	ggtcgtaatc	ccacacactg	gccaatgccg	6480

-continued

---

ccggccctgc	ggaaacctct	acgtgcccgt	ctggaagctc	gtagcggatc	acctcgccag	6540
ctcgtcggtc	acgcttcgac	agacggaaaa	cggccacgtc	catgatgctg	cgactatcgc	6600
gggtgcccac	gtcatagagc	atcggaacga	aaaaatctgg	ttgctcgtcg	cccttgggcg	6660
gcttcctaata	cgacggcgca	ccggctgccc	gcggttgccc	ggattctttg	cggattcgat	6720
cagcggccgc	ttgccacgat	tcaccggggc	gtgcttctgc	ctcgatgcgt	tgcgctggg	6780
cggcctgcgc	ggccttcaac	ttctccacca	ggcctcacc	cagcggccgc	cggatttgta	6840
ccgggcccga	tggtttgoga	ccgctcacgc	cgattcctcg	ggcttggggg	ttccagtgcc	6900
attgcagggc	cggcagacaa	cccagccgct	tacgcctggc	caaccgccc	ttcctccaca	6960
catggggcat	tccacggcgt	cggtgccctg	ttgttcttga	ttttccatgc	cgctccttt	7020
agccgctaaa	attcatctac	tcatttattc	atttgcctat	ttactctggt	agctgcgcga	7080
tgtattcaga	tagcagctcg	gtaatggtct	tgccctggcg	taccgcgtac	atcttcagct	7140
tggtgtgac	ctccgcccgc	aactgaaagt	tgaccgctt	catggctggc	gtgtctgcca	7200
ggctggccaa	cgttgccagc	ttgctgctgc	gtgcgctcgg	acggcccgc	cttagcgtgt	7260
ttgtgctttt	gctcattttc	tctttacctc	attaactcaa	atgagttttg	atttaatttc	7320
agcggccagc	gcctggacct	cgcgggcagc	gtcgcctcgc	ggttctgatt	caagaacggt	7380
tgtgccggcg	gcggcagctg	ctgggtagct	cacgcgctgc	gtgatacggg	actcaagaat	7440
gggcagctcg	taccggccca	gcgcctcggc	aacctcaccc	ccgatgcgcg	tgctttgat	7500
cgcgcgcgac	acgacaaaag	ccgctttag	ccttccatcc	gtgacctcaa	tgcgctgctt	7560
aaccagctcc	accaggtcgg	cggtgcccca	tatgtcgtaa	gggcttggtc	gcaccggaat	7620
cagcacgaag	tcggctgcct	tgatcgcgga	cacagccaa	tccgcccct	ggggcgctcc	7680
gtcgatacct	acgaagtgcg	gcccggccgat	ggccttcacg	tcgcgggtcaa	tcgtcgggcg	7740
gtcgatgcgc	acaacggtta	gcggttgatc	ttcccgcacg	gcccaccaat	cgcgggcact	7800
gcccctgggga	tcggaatcga	ctaacagaa	atcggcccgc	gcgagttgca	gggcgcgggc	7860
tagatggggt	gcgatgctg	tcttgccctg	cccgcctttc	tggttaagta	cagcgataac	7920
ttcatgcgtt	cccttgcgta	ttgtttatt	tactcatcgc	atcatatacg	cagcgaccgc	7980
atgacgcaag	ctgttttact	caaatacaca	tcacctttt	agacggcggc	gctcggtttc	8040
ttcagcggcc	aagctggccg	gccaggccgc	cagcttgcca	tcagacaaa	cggccaggat	8100
ttcatgcagc	cgcacggttg	agacgtgcgc	gggcggctcg	aacacgtacc	cggcccgcgat	8160
catctccgcc	tcgatctcct	cggtaaatgaa	aaacggttcg	tcctggccgt	cctgggtcgg	8220
tttcatgctt	gttctctttg	gcgttcattc	tcggcggccg	ccagggcgctc	ggcctcggtc	8280
aatgcgtcct	cacggaaggc	accgcgccgc	ctggcctcgg	tgggcgtcac	ttcctcgctg	8340
cgctcaagt	cgcggtacag	ggctcagcga	tgcaagccaa	gcagtgcagc	cgctcctttc	8400
acggctcggc	cttctcggtc	gatcagctcg	cgggcgtcgc	cgatctgtgc	cggggtgagg	8460
gtaggcgggg	ggccaaactt	cacgcctcgg	gccttggcgg	cctcgcgccc	gctccgggtg	8520
cggctcgatg	ttagggaaacg	ctcgaactcg	gcaatgcccg	cgaacacggt	caacaccatg	8580
cggccggccg	gcgtggtggt	gtcggcccac	ggctctgcca	ggctacgcag	gcccgcgccc	8640
gcctcctgga	tgcctcggc	aatgtccagt	aggtegcggg	tgctgcgggc	caggcggctc	8700
agcctggtca	ctgtcacaac	gtcgcaggg	cgtagggtggt	caagcatcct	ggccagctcc	8760
gggcggctgc	gcctggtgcc	ggtgatcttc	tcggaaaaca	gcttgggtgca	gcccggccgcg	8820
tgcagttcgg	cccgttggtt	ggtcaagtcc	tggtcgtcgc	tgctgacgcg	ggcatagccc	8880

-continued

---

agcaggccag	cgcgggcgct	cttgttcatg	gcgtaatgtc	tccggttcta	gtcgcaagta	8940
ttctacttta	tgcgactaaa	acacgcgaca	agaaaacgcc	aggaaaaggg	cagggcggca	9000
gcctgtcgcg	taacttagga	cttggtcgac	atgtcgtttt	cagaagacgg	ctgcaactgaa	9060
cgtcagaagc	cgactgcact	atagcagcgg	aggggttggg	ccacaggacg	ggtgtggtcg	9120
ccatgatcgc	gtagtcgata	gtggctccaa	gtagcgaagc	gagcaggact	gggcggcggc	9180
caaagcggtc	ggacagtgtc	ccgagaacgg	gtgcgcatag	aaattgcatc	aacgcatata	9240
gcgctagcag	cacgccatag	tgactggcga	tgctgtcggg	atggacgata	tcccgcgaaga	9300
ggcccggcag	taccggcata	accaagccta	tgctacagc	atccagggtg	acggtgccga	9360
ggatgacgat	gagcgcattg	ttagatttca	tacacgggtc	ctgactgcgt	tagcaattta	9420
actgtgataa	actaccgcat	taaagctagc	ttgcttggtc	gttccgcgtg	aacgtcggtc	9480
cgattgtaoc	tgcgttcaaa	tactttgcca	tcgtgttcgg	cgctgcccg	gtgcgtcggc	9540
tgatctcacg	gatcgactgc	ttctctcgca	acgccatccg	acggatgatg	tttaaaagtc	9600
ccatgtggat	cactccgttg	ccccgtcgtc	caccgtgttg	gggggaaggt	gcacatggct	9660
cagttctcaa	tggaaattat	ctgcctaacc	ggctcagttc	tgcgtagaaa	ccaacatgca	9720
agctccaccg	ggtgcaaagc	ggcagcggcg	gcaggatata	ttcaattgta	aatggcttca	9780
tgccgggaa	atctacatgg	atcagcaatg	agtatgatgg	tcaatatgga	gaaaaagaaa	9840
gagtaattac	caattttttt	tcaattcaaa	aatgtagatg	tccgcagcgt	tattataaaa	9900
tgaaagtaca	ttttgataaa	acgacaaaat	acgatccgtc	gtatttatag	gcgaaagcaa	9960
taaacaaaat	attctaattc	ggaaatcttt	atctcgacgt	gtctacattc	acgtccaaat	10020
gggggcttag	atgagaaact	tcacgatcga	tgcttgattg	tcgccattcc	cagataccca	10080
tttcatcttc	agattgggtc	gagattatgc	gaaaatatac	actcatatac	ataaatactg	10140
acagtttgag	ctaccaattc	agtgtagccc	attacctcac	ataattcact	caaatgctag	10200
gcagtctgtc	aactcggcgt	caatttgctg	gccactatac	gatagttgcg	caaattttca	10260
aagtccctggc	ctaacatcac	acctctgtcg	gcggcgggtc	ccattttgtg	taaatccacc	10320
atatcgaatt	aattcagact	cctttgcccc	agagatcaca	atggacgact	tcctctatct	10380
ctacgatcta	gtcaggaagt	tcgacggaga	aggtgacgat	accatgttca	ccactgataa	10440
tgagaagatt	agccttttca	atctcagaaa	gaatgctaac	ccacagatgg	ttagagaggc	10500
ttacgcagca	ggctctcatc	agacgatcta	cccagcaat	aatctccagg	agatcaaata	10560
ccttcccaag	aagggttaaag	atgcagtcaa	aagattcagg	actaactgca	tcaagaacac	10620
agagaaagat	atatttttca	agatcagaag	tactattcca	gtatggacga	ttcaaggctt	10680
gcttcacaaa	ccaaggcaag	taatagagat	tggagtctct	aaaaaggtag	ttcccactga	10740
atcaaaggcc	atggagtcaa	agattcaaat	agaggacctc	acagaactcg	ccgtaaagac	10800
tggcgaacag	ttcatacaga	gtctcttacg	actcaatgac	aagaagaaaa	tcttcgtaaa	10860
catggtggag	cacgacacgc	ttgtctactc	caaaaatatac	aaagatacag	tctcagaaga	10920
ccaaagggca	attgagactt	ttcaacaaaag	ggtaatatcc	ggaaacctcc	tcggattcca	10980
ttgccagct	atctgtcact	ttattgtgaa	gatagtgtaa	aaggaagggtg	gctcctacaa	11040
atgccatcat	tgcgataaag	gaaaggccat	cgttgaagat	gcctctgccg	acagtggctc	11100
caaagatgga	ccccaccoca	cgaggagcat	cgtggaaaaa	gaagacgttc	caaccacgtc	11160
ttcaaagcaa	gtggattgat	gtgatctctc	cactgacgta	agggatgacg	cacaatccca	11220

-continued

---

ctatccttcg caagaccctt cctctatata aggaagtcca ttccatttgg agaggacacg	11280
ctgaaatcac cagtctccaa gcttgccggg atcgtttcgc atgattgaac aagatggatt	11340
gcacgcaggt tctccggcgc cttgggtgga gaggtattc ggctatgact gggcacaaca	11400
gacaatcggc tgctctgatg ccgccgtgtt ccggctgtca gcgcaggggc gcccggttct	11460
ttttgtcaag accgacctgt ccgggtgccct gaatgaactg caggacgagg cagcgcggct	11520
atcgtggctg gccacgacgg gcgttccttg cgcagctgtg ctgcagcttg tcaactgaagc	11580
gggaagggac tggctgctat tgggcgaagt gccggggcag gatctcctgt catctcacct	11640
tgctcctgcc gagaagtat ccatcatggc tgatgcaatg cggcggctgc atacgcttga	11700
tccggctacc tgcccattcg accaccaagc gaaacatcgc atcgagcggc cactgactcg	11760
gatggaagcc ggtcttctg atcaggatga tctggacgaa gagcatcagg ggctcgcgcc	11820
agccgaactg ttcgccaggc tcaaggcgcg catgcccgac ggcgaggatc tcgtcgtgac	11880
ccatggcgat gcctgcttgc cgaatatcat ggtggaaaat ggccgctttt ctggattcat	11940
cgactgtggc cggctgggtg tggcggaccg ctatcaggac atagcgttgg ctaccctgga	12000
tattgctgaa gagcttggcg gcgaatgggc tgaccgcttc ctctgtcttt acggtatcgc	12060
cgctcccgat tcgcagcgca tcgccttcta tcgccttctt gacgagtctt tctgagcggg	12120
actctggggt tcgaaatgac cgaccaagcg acgcccacc tgccatcacg agatttcgat	12180
tccaccgccc ccttctatga aaggttgggc ttcggaatcg tttccggga cgcgcgtgg	12240
atgatcctcc agcgcgggga tctcatgctg gagttcttcg cccaccocgg atcgatccaa	12300
cacttacgtt tgcaacgtcc aagagcaaat agaccacgaa cgcgggaagg ttgccgcagc	12360
gtgtggattg cgtctcaatt ctctcttgca ggaatgcaat gatgaatatg atactgacta	12420
tgaaactttg agggaatact gcttagcacc gtcacctcat aacgtgcac atgcatgccc	12480
tgacaacatg gaacatcgtc atttttctga agaattatgc tcgttgagg atgtcgcggc	12540
aattgcagct attgccaaca tcgaaactacc cctcacgcat gcattcatca atattattca	12600
tgccgggaaa ggcaagatta atccaactgg caaatcatcc agcgtgattg gtaacttcag	12660
ttccagcgac ttgattcgtt ttggtgctac ccacgttttc aataaggacg agatggtgga	12720
gtaaagaagg agtgcgctga agcagatcgt tcaaactttt ggcaataaag tttcttaaga	12780
ttgaatcctg ttgccggtct tgccgatgatt atcatataat ttctggtgaa ttacgttaag	12840
catgtaataa ttaacatgta atgcatgacg ttatttatga gatgggtttt tatgattaga	12900
gtcccgcaat tatacattta atacgcgata gaaaacaaaa tatagcgcgc aaactaggat	12960
aaattatcgc gcgcgggtgc atctatgtta ctagatcgat caaacttcgg tactgtgtaa	13020
tgacgatgag caatcgagag gctgactaac aaaaggtaca tcgcgatgga tcgatccatt	13080
cgccattcag gctcgcgaac tgttgggaag ggcgatcggg gcgggcctct tcgctattac	13140
gccagctggc gaaaggggga tgtgctgcaa ggcgattaag ttgggtaacg ccagggtttt	13200
cccagtcacg acgttgtaaa acgacggcca gtgaattcct gcagcccggg ggatccgccc	13260
actcgaggcg cgcctgcgac ggatataatg agccgtaaac aaagatgatt aagtagtaat	13320
taatacgtac tagtaaaagt ggcaaaagat aacgagaaag aaccaatttc tttgcattcg	13380
gccttagcgg aaggcatata taagctttga ttattttatt tagtgtaatg atttcgtaca	13440
accaaagcat ttatttagta ctctcacact tgtgtcgcgg ccggccgcta caggaaacag	13500
tggtgccggc cctcggcgcg ctctgactgc tccacgatgg tgtagtcctc gttgtgggag	13560
gtgatgtcca gcttgagtc cacgtagtag tagccgggca gctgcacggg cttcttggcc	13620

-continued

---

atgtagatgg	acttgaactc	caccaggtag	tggcgcgctg	ccttcagctt	cagggccttg	13680
tggatctcgc	ccttcagcac	gccgtcgcgg	gggtacaggc	gctcggtgga	ggcctcccag	13740
cccatagtct	tcttctgcat	tacggggccg	tcggagggga	agttcacgcc	gatgaacttc	13800
accttgtaga	tgaaggagcc	gtcctgcagg	gaggagtctc	gggtcacggt	caccacgccc	13860
ccgtcctcga	agttcatcac	gcgctcccac	ttgaagcctc	cggggaagga	cagcttcttg	13920
tagtcgggga	tgtcggcggg	gtgcttcacg	tacaccttgg	agccgtaactg	gaactggggg	13980
gacaggatgt	cccaggcgaa	gggcaggggg	ccgcccttgg	tcaccttcag	cttggcggtc	14040
tgggtgcctc	cgtagggggc	gccctcgcgc	tcgccctcga	tctcgaactc	gtggccggtc	14100
acggagccct	ccatgcgcac	cttgaagcgc	atgaactcct	tgatgacgtc	ctcggaggag	14160
gccatggggc	gcttgggggg	ctatggaaga	ctttcttagt	tagttgtgtg	aataagcaat	14220
gttgggagaa	tcgggactac	ttataggata	ggaataaaac	agaaaagtat	taagtgctaa	14280
tgaaatattt	agactgataa	ttaaaatctt	cacgtatgtc	cacttgatat	aaaaacgtca	14340
ggaataaagg	aagtacagta	gaattttaaag	gtactctttt	tatatatacc	cgtgttctct	14400
ttttggctag	ctagttgcat	aaaaaataat	ctatatTTTT	atcattattt	taaatactct	14460
atgagatggg	aaatatttat	cataattttt	tttactatta	tttattattt	gtgtgtgtaa	14520
tacatataga	agtttaattac	aaattttatt	tactttttca	ttattttgat	atgattcacc	14580
attaatttag	tgttattatt	tataatagtt	cattttaatc	ttttgtata	tattatgctg	14640
gcagtacttt	tttctacat	ataactacta	ttacatttta	tttatataat	atttttatta	14700
atgaattttc	gtgataatat	gtaatatgtg	tcattattat	ttcagatttt	ttaaaaatat	14760
ttgtgttatt	atztatgaaa	tatgtaattt	ttttagattt	tgattttatg	atgataaagt	14820
gttctaaatt	caaaaagagg	gggaaagcgt	aaacattaaa	aaacgctatc	aaacaaaaac	14880
aaaatcttgt	taataaagat	aaaactggtt	gttttgatca	ctgttatttc	gtaataataa	14940
aacattattt	atatttatat	tgttgacaac	caaatttgcc	tatcaaatct	aaccaatata	15000
atgcatgcgt	ggcaggtaat	gtactaccat	gaacttaagt	catgacataa	taaaccgtga	15060
atctgaccaa	tgcattgacc	tanctaaatt	gtattttgta	cacgaagcaa	atgattcaat	15120
tcacaatgga	gatgggaaac	aaataatgaa	gaaccocagaa	ctaagaaagc	tttctgaaa	15180
aataaaataa	aggcaatgtc	aaaagtatac	tgcacatcca	gtccagaaag	cacatgatat	15240
ttttttatca	gtatcaatgc	agctagtttt	attttacaat	atcgatatag	ctagtttaaa	15300
tatattgcag	ctagatttat	aaatatttgt	gttattattt	atcatttggtg	taatcctggt	15360
tttagtattt	tagtttatat	atgatgataa	tgtattccaa	atttaaaaga	agggaaataa	15420
atttaaacaa	gaaaaaaagt	catcaaacaa	aaaacaaatg	aaaggggtgga	aagatgttac	15480
catgtaatgt	gaatgttaca	gtatttcttt	tattatagag	ttaacaaatt	aactaatatg	15540
attttggtaa	taatgataaa	atattttttt	tattattatt	tcataatata	aaaatagttt	15600
acttaataata	aaaaaaatc	tatcgttcac	aacaaagtgt	gccacctaat	ttaaccatgc	15660
atgtacccat	ggaccatatt	aggtaacccat	caaacctgat	gaagagataa	agagatgaag	15720
acttaagtca	taacacaaaa	ccataaaaaa	caaaaataca	atcaaccgctc	aatctgacca	15780
atgcatgaaa	aagctgcaat	agtgagtggc	gacacaaagc	acatgatattt	cttacaacgg	15840
agataaaaacc	aaaaaaatat	ttcatgaaca	acctagaaca	aataaagctt	ttatataata	15900
aatatataaa	taaataaagg	ctatggaata	atatacttca	atatatttgg	attaaataaa	15960

-continued

---

ttggtggcgg	ggttgatata	tttatacaca	cctaaagtca	cttcaatctc	atthtcaett	16020
aacttttatt	ttttttttct	ttttatttat	cataaagaga	atattgataa	tatacttttt	16080
aacatatttt	tatgacattt	tttattggtg	aaaacttatt	aaaaatcata	aatthtgtaa	16140
gtagattta	tttaaagagt	tcctcttctt	atthttaaatt	ttttaaataa	ttthtaata	16200
actaaattt	gtgttaaaaa	tggttaaaaa	gtgtgttatt	aaccttctc	ttcgaggatc	16260
cgtacgatcc	cacatgcaag	ttttatttcc	aatccctttt	cctttgaata	actgaccaag	16320
aacaacaaga	aaaaaaaaaa	aaaagaaaag	gatcattttg	aaaggatatt	ttcgcctcct	16380
attcaaatac	tgtattttta	ccaaaaaac	tgtatttttc	ctacactctc	aagctttggt	16440
tttcgcttcg	actctcatga	tttccttcat	atgccaatca	ctctatttat	aaatggcata	16500
aggtagtgtg	aacaattgca	aagcttgca	tcaaaagctt	gcaatgtaca	aattaatggt	16560
tttcatgcct	ttcaaaatta	tctgcacccc	ctagctatta	atctaacatc	taagtaaggc	16620
tagtgaatth	tttgaatag	tcatgcagtg	cattaatttc	cccgtagacta	ttttggcttt	16680
gactccaaca	ctggcccctg	acatccgtcc	ctcattacat	gaaaagaaat	attgthtata	16740
ttcttaatta	aaaatattgt	cccttctaaa	ttttcatata	gttaattatt	atattacttt	16800
tttctctatt	ctattagttc	tattttcaaa	ttattattta	tgcatatgta	aagtacatta	16860
tatttttgct	atatacttaa	atatttctaa	attattaaaa	aaagactgat	atgaaaaatt	16920
tattcttttt	aaagctatat	cattttatat	atactttttc	ttttcttttc	tttcattttc	16980
tattcaatth	aataagaaat	aaatthtgta	aaatthtatt	tatcaattta	taaaaatatt	17040
ttactttata	tgthttttca	cattthtggt	aaacaaatca	tatcattatg	attgaaagag	17100
aggaaattga	cagtgtgtaa	taagtgtgta	gaaaaaaatg	tgthatttcc	taaaaaaac	17160
ctaaacaaac	atgtatctac	tctctatttc	atctatctct	catttcattt	ttctctttat	17220
ctctttcttt	atthttttat	catatcattt	cacattaatt	atthttactc	tctttatttt	17280
ttctctctat	ccctctctta	tttccactca	tatatacact	ccaaaaattgg	ggcatgcctt	17340
tatcactact	ctatctctc	cactaaatca	tttaaatgaa	actgaaaagc	attggcaagt	17400
ctctccctct	cctcaagtga	tttccaactc	agcattggca	tctaattgat	tcagtatctc	17460
tattgcatgt	gtaaaagtct	ttccacaata	cataactatt	aattaatctt	aaataaataa	17520
aggataaaat	atthtttttt	cttcataaaa	ttaaaatag	ttatthtttg	tttagatgta	17580
tattcgaata	aatctaaata	tatgataatg	atthttttata	ttgattaaac	atataatcaa	17640
tattaatat	gatatttttt	tatatagggt	gtacacataa	ttttataagg	ataaaaaata	17700
tgataaaaaat	aaatthttaa	tattthttata	tttacgagaa	aaaaaaatat	tttagccata	17760
aataaatgac	cagcatattt	tacaacctta	gtaattcata	aattcctata	tgtatatttg	17820
aaatthaaaa	cagataatcg	ttaaaggaag	gaatcctacg	tcctctcttg	ccattthggt	17880
ttcatgcaaa	cagaaagga	cgaaaaacca	cctcaacctg	aatcactctt	cacaccattt	17940
ttactagcaa	acaagtctca	acaactgaag	ccagctctct	ttccgthtct	ttttacaaca	18000
ctthctttga	aatagtagta	ttthtttttc	acatgattta	ttaacgtgcc	aaaagatgct	18060
tattgaatag	agtgcacatt	tgtaatgtac	tactaattag	aacatgaaaa	agcattthgtc	18120
taacacgata	atcctgtgaa	ggcgttaact	ccaaagatcc	aatthcacta	tataaattgt	18180
gacgaaagca	aaatgaaatc	acatagctga	gagagaaagg	aaaggttaac	taagaagcaa	18240
tacttcagcg	gccgatgga	gagatctcaa	cggcagtctc	ctccgccacc	gtcgcctgctc	18300
tcctctctgt	cctccgtctc	cgcggacacc	gtcctctctc	ctcccgaaa	gagcgggag	18360

-continued

---

```

gcggcgacgg ccaaggccgg cgccgagcct aataagagga tccgcaagga ccccgccgc 18420
gccgcccggg ggaagaggag ctccgtctac aggggagtc aagggcacag gtggacgggc 18480
aggttcgagg cgcatctctg ggacaagcac tgcctcgccg cgctccacaa caagaagaaa 18540
ggcaggcaag tctacctggg ggcgtatgac agcgaggagg cagctgctcg tgcctatgac 18600
ctcgcagctc tcaagtaactg gggctctgag actctgctca acttcctgt ggaggattac 18660
tccagcgaga tgcgggagat ggaggccgtg tcccgggagg agtacctggc ctccctccgc 18720
cgcaggagca gcgctctctc cagggcgctc tccaagtaca gaggcgtcgc caggcatcac 18780
cacaacggga ggtgggaggc acggattggg cgagtctttg ggaacaagta cctctacttg 18840
ggaacatttg aactcaaga agaggcagcc aaggcctatg accttgccgc cattgaatac 18900
cgtggcgtca atgctgtaac caacttcgac atcagctgct acctggacca cccgctgttc 18960
ctggcacagc tccaacagga gccacaggtg gtgccggcac tcaaccaaga acctcaacct 19020
gatcagagcg aaaccggaac tacagagcaa gagccggagt caagcgaagc caagacaccg 19080
gatggcagtg cagaaccoga tgagaacgag gtgctgagc acaccgcgga gccctcacc 19140
acagtcgacg acagcatcga agagggttg tggagccctt gcatggatta cgagctagac 19200
accatgtcga gaccaaactt tggcagctca atcaatctga gcgagtgggt cgctgacgca 19260
gacttcgact gcaacatcgg atgcctgttc gatgggtggt ctgcccgtga cgaaggaagc 19320
aaggatggtg taggtctggc agatttcagt ctggttgagg caggtgatgt ccagctgaag 19380
gatgttcttt cggatatgga agaggggata caacctccag cgatgatcag tgtgtgcaac 19440
gcgcccgcat gagccgtaaa ggttcaatac aacgagtgtc tgttttctta gggacaagca 19500
ttgtacttat gtatgattct gtgtaacct gagtcttcca cgttgacta atgtgaaggg 19560
caaaaataaa acacagaaca agttcgtttt tctcaataa tgtgaaggta gaaaatggaa 19620
ccatgcctcc tctcttgcac gtgatttaaa atattagcag atggtaccgt acgagatccg 19680
gccggccaga tcctgcagga gatccaagct tgg 19713

```

```

<210> SEQ ID NO 33
<211> LENGTH: 10287
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: vector pKR1144
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: (7136)..(7136)
<223> OTHER INFORMATION: n is a, c, g, or t
<400> SEQUENCE: 33

```

```

gtacgagatc cggccggcca gatcctgcag gagatccaag cttggcgcgc cgttctatag 60
tgtcacctaa atcgtatgtg tatgatacat aaggttatgt attaattgta gccgcgttct 120
aacgacaata tgtccatgat gtgcactctc agtacaatct gctctgatgc ccatagtta 180
agccagcccc gacacccgcc aacaccgct gacgcgccct gacgggcttg tctgctcccg 240
gcatccgctt acagacaagc tgtgaccgtc tccgggagct gcatgtgtca gaggttttca 300
ccgtcatcac cgaaacgcgc gagacgaaag ggcctcgtga tacgcctatt tttataggtt 360
aatgtcatga ccaaaatccc ttaacgtgag ttttcgttcc actgagcgtc agaccccgta 420
gaaaagatca aaggatcttc ttgagatcct tttttctg cgcgtaatctg ctgcttgcaa 480
acaaaaaac caccgctacc agcgggtggt tgtttgccc atcaagagct accaactctt 540

```

-continued

---

tttccgaagg taactggctt cagcagagcg cagatacca atactgtcct tctagtgtag	600
ccgtagttag gccaccactt caagaactct gtagcaccgc ctacatacct cgctctgcta	660
atcctgttac cagtggctgc tgcagtggc gataagtcgt gtcttacogg gttggactca	720
agacgatagt taccggataa ggcgcagcgg tcgggctgaa cggggggttc gtgcacacag	780
cccagcttgg agcgaacgac ctacaccgaa ctgagatacc tacagcgtga gcattgagaa	840
agcgcaccgc ttcccgaagg gagaaaggcg gacaggtatc cggtaaagcgg cagggtcgga	900
acaggagagc gcacgagggg gcttccaggg ggaaacgcct ggtatcttta tagtctgtc	960
gggtttcgcc acctctgact tgagcgtcga tttttgtgat gctcgtcagg ggggcgagc	1020
ctatgaaaa acgccagcaa cgcggccttt ttacggttcc tggccttttg ctggcctttt	1080
gctcacatgt tctttctgc gttatccct gattctgtgg ataaccgtat taccgccttt	1140
gagtgagctg ataccgctcg ccgcagccga acgaccgagc gcagcagtc agtgagcag	1200
gaagcggaa agcgcaccaat acgcaaaccg cctctcccc cgcggtggcc gattcattaa	1260
tgcaggttga tcgattcgac atcgatctag taacatagat gacaccgcgc gcgataattt	1320
atcctagttt gcgcgctata ttttgtttc tatcgcgtat taaatgtata attgcgggac	1380
tctaatacata aaaaccatc tcataataa cgtcatgcat tacatgttaa ttattacatg	1440
cttaacgtaa ttcaacagaa attatatgat aatcatcgca agaccggcaa caggattcaa	1500
tcttaaaaa ctttattgcc aaatgtttga acgatctgct tcgacgcact ccttctttag	1560
gtacctcact attcctttgc cctcggacga gtgctggggc gtcggtttcc actatcggcg	1620
agtacttcta cacagccatc ggtccagacg gccgcgcttc tgcggcgat ttgtgtacgc	1680
ccgacagtcc cggctccgga tcggacgatt gcgctgcac gacctgcgc ccaagctgca	1740
tcacgaaat tgcgctcaac caagctctga tagagtgggt caagaccaat gcggagcata	1800
tacgcccgga gccgcggcga tcttgcaagc tccggatgcc tccgctcga gtagcgcgtc	1860
tgctgctcca tacaagccaa ccacggcctc cagaagaaga tgttggcgac ctcgatttg	1920
gaatccccga acatcgctc gctccagtca atgaccgctg ttatgcggcc attgtccgtc	1980
aggacattgt tggagccgaa atccgcgtgc acgaggtgcc ggacttcggg gcagtcctcg	2040
gccaaagca tcagctcacc gagagcctgc gcgacggacg cactgacggt gtcgtccatc	2100
acagtttgcc agtgatacac atggggatca gcaatcgcgc atatgaaatc acgccatgta	2160
gtgtattgac cgattccttg cgttccgaat gggccgaacc cgctcgtctg gctaagatcg	2220
gccgcagcga tcgcatccat ggctcccgcg accggctgca gaacagcggg cagttcggtt	2280
tcaggcaggt cttgcaacgt gacaccctgt gcacggcggg agatgcaata ggtcaggctc	2340
tcgctgaatt cccaatgct aagcacttcc ggaatcggga gcgcggccga tgcaaatgct	2400
cgataaacat aacgatcttt gtagaaacca tcggcgcagc tatttaccgg caggacatat	2460
ccacgcctc ctacatcgaa gctgaaagca cgagattctt cgcctccga gagctgcac	2520
aggtcggaga cgctgtcgaa cttttcgatc agaaacttct cgacagacgt cgcggtgagt	2580
tcaggctttt tcactggtta ataagaagag aaaagagttc ttttgttatg gctgaagtaa	2640
tagagaaatg agctcgagcg tgcctctcc aaatgaaatg aacttcctta tatagaggaa	2700
gggtcttgcg aaggatagtg ggattgtgct tcatecctta cgtaagtgga gatgtcacat	2760
caatecaact gctttgaaga cgtggttga acgtcttctt tttccacgat gctcctcgtg	2820
ggtgggggtc catctttggg accactgtcg gcagaggcat cttgaatgat agcctttcct	2880
ttatcgcaat gatggcattt gtaggagcca ccttctttt ctactgtcct ttcgatgaag	2940

-continued

---

tgacagatag	ctgggcaatg	gaatccgagg	aggtttcccg	aaattatcct	ttgttgaaaa	3000
gtctcaatag	ccctttggtc	ttctgagact	gtatctttga	catttttggg	gtagaccaga	3060
gtgtcgtgct	ccaccatggt	gacgaagatt	ttcttcttgt	cattgagtcg	taaaagactc	3120
tgtatgaact	gttcgccagt	cttcacggcg	agttctgtta	gatcctcgat	ttgaatctta	3180
gactccatgc	atggccttag	attcagtagg	aactaccttt	ttagagactc	caatctctat	3240
tacttgccctt	ggtttatgaa	gcaagccttg	aatcgtccat	actggaatag	tacttctgat	3300
cttgagaaat	atgtctttct	ctgtgttctt	gatgcaatta	gtcctgaate	ttttgactgc	3360
atctttaacc	ttcttgggaa	ggtatttgat	ctcctggaga	ttgttactcg	ggtagatcgt	3420
cttgatgaga	cctgctcgtg	aggcctctct	aaccatctgt	gggtcagcat	tctttctgaa	3480
attgaagagg	ctaaccttct	cattatcagt	ggtgaacata	gtgtcgtcac	cttcaccttc	3540
gaacttcctt	cctagatcgt	aaagatagag	gaaatcgtcc	attgtaatct	ccggggcaaaa	3600
ggagatctct	tttggggctg	gatcactgct	gggccttttg	gttcctagcg	tgagccagtg	3660
ggctttttgc	tttgggtggc	ttgttagggc	cttagcaaaag	ctcttgggct	tgagttgagc	3720
ttctcctttg	gggatgaagt	tcaacctgtc	tgtttgctga	cttgttgtgt	acgcgtcagc	3780
tgctgctctt	gcctctgtaa	tagtggcaaaa	tttcttgtgt	gcaactccgg	gaacgcgctt	3840
tgttgccgcc	tttgtaacaac	cccagtcate	gtatataccg	gcatgtggac	cgttatacac	3900
aacgtagtag	ttgatatgag	ggtgttgaat	acccgattct	gctctgagag	gagcaactgt	3960
gctgttaagc	tcagattttt	gtgggattgg	aattggatcg	atctogatcc	cgcgaaatta	4020
atagactca	ctatagggag	accacaacgg	tttccctcta	gaaataattt	tgtttaactt	4080
taagaaggag	atatacccat	ggaaaagcct	gaactcaccg	cgacgtctgt	cgagaagttt	4140
ctgatcgaaa	agttcgacag	cgtctccgac	ctgatgcagc	tctcggaggg	cgaagaatct	4200
cgtgctttca	gcttcgatgt	aggagggcgt	ggatatgtcc	tgcgggtaaa	tagctgcgcc	4260
gatggtttct	acaaagatcg	ttatgtttat	cggcactttg	catcggccgc	gctcccattt	4320
ccggaagtgc	ttgacattgg	ggaatcagc	gagagcctga	cctattgcat	ctcccgcctt	4380
gcacaggggtg	tcacgttgca	agacctgcct	gaaaccgaac	tgcocgctgt	tctgcagccg	4440
gtcgcggagg	ctatggatgc	gatcgtctcg	gccgatctta	gccagacgag	cggttccggc	4500
ccattcggac	cgcaaggaat	cgttcaatac	actacatggc	gtgatttcat	atgcgcgatt	4560
gctgatcccc	atgtgatca	ctggcaaaact	gtgatggacg	acaccgtcag	tgcgtccgtc	4620
gcgcaggctc	tcgatgagct	gatgctttgg	gccgaggact	gcccogaagt	ccggcaactc	4680
gtgcacgcgg	atttcggctc	caacaatgtc	ctgacggaca	atggccgcat	aacagcggtc	4740
attgactgga	gcgaggcgtg	gttcggggat	tcccaatacg	aggtcgccaa	catcttcttc	4800
tggaggccgt	ggttggcttg	tatggagcag	cagacgcgct	acttcgagcg	gaggcatccg	4860
gagcttgacg	gatcgcgcgg	gctccggggc	tatatgtctc	gcattggtct	tgaccaactc	4920
tatcagagct	tggttgacgg	caatttcgat	gatgcagctt	gggcgcaggg	tcgatgcgac	4980
gcaatcgtcc	gatccggagc	cgggactgtc	gggcgtacac	aaatcgcccc	cagaagcgcg	5040
gccgtctgga	ccgatggctg	tgtagaagta	ctcgcggata	gtggaaaccg	acgccccagc	5100
actcgtccga	gggcaaaagga	atagttaggt	acagcttggg	tcgatccggc	tgctaacaaa	5160
gcccgaaggg	aagctgagtt	ggctgctgcc	accgctgagc	aataactagc	ataacccttt	5220
ggggcctcta	aacgggtcct	gaggggtttt	ttgctgaaag	gaggaactat	atccggatga	5280

-continued

---

tcgggcgcgc	cgtcgacgga	tccactagtt	ctagagcggc	ccgcgcgcgc	gacggatata	5340
atgagccgta	aacaagatg	attaagtagt	aattaatacg	tactagtaaa	agtggaacaa	5400
gataacgaga	aagaaccaat	ttctttgcat	tggccttag	cggaaggcat	atataagctt	5460
tgattatfff	atftagtgta	atgatttcgt	acaaccaaag	catttatftha	gtactctcac	5520
acttggtgco	cggcgggcoo	ctacaggaac	agggtggggc	ggccctcggc	gcgctcgtac	5580
tgctccacga	tggtgtagtc	ctcgttgtag	gaggtagatg	ccagcttgga	gtccacgtag	5640
tagtagccgg	gcagctgcac	gggcttcttg	gccatgtaga	tggacttgaa	ctccaccagg	5700
tagtggccgc	cgctctcag	cttcagggcc	ttgtggatct	cgccctcag	cacgcgctcg	5760
cgggggtaca	ggcgcctcgg	ggaggcctcc	cagcccatag	tcttctctcg	cattacgggg	5820
ccgtcggagg	ggaagtccac	gccgatgaac	ttcaccttgt	agatgaagga	gccgtcctgc	5880
agggaggagt	cctgggtcac	ggtcaccacg	ccgcctcct	cgaagttcat	cacgcgctcc	5940
cacttgaagc	cctcggggaa	ggacagcttc	ttgtagtcgg	ggatgtcggc	gggggtgcttc	6000
acgtacacct	tggagccgta	ctggaactgg	ggggacagga	tgtcccaggc	gaagggcagg	6060
gggcccctc	tggtcacctt	cagcttgggc	gtctgggtgc	cctcgtaggg	gcggccctcg	6120
ccctcgcct	cgatctcgaa	ctcgtggcgg	ttcacggagc	cctccatgcg	caccttgaag	6180
cgcataaact	ccttgatgac	gtcctcggag	gagggccatg	gccgcttggg	gggctatgga	6240
agactttctt	agttagttgt	gtgaataaag	aatgttggga	gaatcgggac	tacttatag	6300
ataggaataa	aacagaaaag	tattaagtgc	taatgaaata	tttagactga	taattaaaat	6360
cttcacgtat	gtccacttga	tataaaaacg	tcaggaataa	aggaagtaca	gtagaattta	6420
aaggtactct	ttttatata	accctgttcc	tctttttggc	tagctagtgg	cataaaaaat	6480
aatctatatt	tttatcatta	ttttaaatat	cttatgagat	ggtaaatatt	tatcataatt	6540
ttttttacta	ttatttatta	tttgtgtgtg	taatacatat	agaagttaat	tacaaatfff	6600
atttactfff	tcattatfff	gatatgattc	accattaatt	tagtgttatt	atttataata	6660
gttcatttfa	atctttttgt	atatattatg	cgtgcagtag	ttttttccta	catataacta	6720
ctattacatt	ttatttatat	aatattttta	ttaatgaatt	ttcgtgataa	tatgtaatat	6780
tgttcattat	tatttcagat	tttttaaaaa	tatttgtgtt	attatftatg	aaatattgta	6840
tttttttagt	atttgatttt	atgatgataa	agtgttctaa	attcaaaaga	aggggggaaag	6900
cgtaaacatt	aaaaaacgtc	atcaaacaaa	aacaaaatct	tgtaataaaa	gataaaaactg	6960
tttgttttga	tcactgttat	ttcgtaatat	aaaaacatta	tttatattta	tattgttgac	7020
aaccaaattt	gcctatcaaa	tctaaccaat	ataatgcatg	cgtggcagg	aatgtactac	7080
catgaactta	agtcatgaca	taataaacgg	tgaatctgac	caatgcatgt	acctanctaa	7140
attgtatttg	tgacacgaag	caaatgattc	aattcacaat	ggagatggga	aacaaataat	7200
gaagaaccca	gaactaagaa	agcttttctg	aaaaataaaa	taaaggcaat	gtcaaaagta	7260
tactgcatca	tcagtccaga	aagcacatga	tattttttta	tcagtatcaa	tgcagctagt	7320
tttattttac	aatatcgata	tagctagttt	aaatatattg	cagctagatt	tataaatatt	7380
tgtgttatta	tttatcattt	gtgtaatcct	gttttttagta	ttttagttta	tatatgatga	7440
taatgtattc	caaatttaaa	agaagggaaa	taaattttaa	caagaaaaaa	agtcatacaa	7500
caaaaaacaa	atgaagggtg	ggaagatgt	taccatgtaa	tgtgaatgtt	acagtatttc	7560
ttttattata	gagttaacaa	attaactaat	atgattttgt	taataatgat	aaaatatttt	7620
ttttattatt	atttcataat	ataaaaaatg	tttacttaat	ataaaaaaaa	ttctatcgtt	7680

-continued

---

cacaacaaag	ttggccaact	aatttaacca	tgcatgtacc	catggaccat	attaggtaac	7740
catcaaacct	gatgaagaga	taaagagatg	aagacttaag	tcataacaca	aaaccataaa	7800
aaacaaaaat	acaatcaacc	gtcaatctga	ccaatgcatg	aaaaagctgc	aatagtgagt	7860
ggcgacacaa	agcacatgat	tttcttacia	cggagataaa	accaaaaaaa	tatttcatga	7920
acaacctaga	acaataaag	cttttatata	ataaatatat	aaataaataa	aggctatgga	7980
ataatatact	tcaatatatt	tggattaaat	aaattgttgg	cggggttgat	atatttatac	8040
acacctaag	tcacttcaat	ctcattttca	cttaactttt	atTTTTTTTT	tctttttatt	8100
tatcataaag	agaatattga	taataactt	tttaacatat	ttttatgaca	ttttttattg	8160
gtgaaaactt	attaaaaatc	ataaattttg	taagttagat	ttatttaaag	agttcctcct	8220
cttattttta	atTTTTTaat	aaatTTTTta	ataactaaaa	tttTgttaa	aaatgttaa	8280
aaagtgtgtt	attaaccott	ctcttcgagg	atccgtaccg	agctcggatc	ctctagaaat	8340
ccgtcaacat	ggTggagcac	gacactctcg	tctactccaa	gaatatcaaa	gatacagtct	8400
cagaagacca	aagggctatt	gagacttttc	aacaaagggT	aatatcgTga	aacctcctcg	8460
gattccattg	cccagctatc	tgTcacttca	tcaaaaggac	agtagaaaag	gaaggtggca	8520
cctacaaatg	ccatcattgc	gataaaggaa	aggctatcgt	tcaagatgcc	tctgcccgaca	8580
gtggtcccaa	agatggaccc	cccccacga	ggagcatcgt	ggaaaaagaa	gacgttccaa	8640
ccacgtcttc	aaagcaagtg	gattgatgtg	atgatcctat	gcgtatggta	tgacgtgtgt	8700
tcaagatgat	gacttcaaac	ctacctatga	cgtatggTat	gaacgtgtgt	cgactgatga	8760
cttagatcca	ctcagagcgc	tataaatagc	tacctacgca	ccctgcgcta	ccatccctag	8820
agctgcagct	tatttttaca	acaattacca	acaacaacaa	acaacaaaca	acattacaat	8880
tactatttac	aattacagtc	gaccggggat	cgtacctcta	gggtggcggc	cgcaagtatg	8940
aactaaaatg	catgtaggTg	taagagctca	tggagagcat	ggaatattgt	atccgaccat	9000
gtaacagtat	aataactgag	ctccatctca	cttcttctat	gaataacaa	aggatgttat	9060
gatataatga	caactctatc	atgcacctta	ttgttctatg	ataaatttcc	tcttattatt	9120
ataaatcatc	tgaatcgtga	cggcttatgg	aatgcttcaa	atagtacaaa	aacaaatgtg	9180
tactataaga	ctttctaaac	aattctaacc	ttagcattgt	gaacgagaca	taagtgttaa	9240
gaagacataa	caattataat	ggaagaagtt	tgtctccatt	tatatattat	atattaccca	9300
cttatgtatt	atattaggat	gttaaggaga	cataacaatt	ataaagagag	aagtttTgat	9360
ccattttat	attatatact	accattttat	atattatact	tatccactta	tttaatgtct	9420
ttataaggtt	tgatccatga	tatttcta	atTTtagttg	atatgtatat	gaaagggTac	9480
tatttgaact	ctcttactct	gtataaaggt	tggatcatcc	ttaaagTggg	tctatttTaat	9540
tttattgctt	cttacagata	aaaaaaaaat	tatgagTtgg	tttgataaaa	tattgaagga	9600
tttaaaataa	taataaataa	catataatat	atgtatataa	atTTattata	atataacatt	9660
tatctataaa	aaagTaaata	ttgtcataaa	tctatacaat	cgttttagcct	tgtcggacga	9720
atctcaatta	tttaaacgag	agTaaacata	tttgactttt	tggTtattta	acaaattatt	9780
atttaacact	atatgaaatt	ttttttttta	tcagcaaaaga	ataaaattaa	attaagaagTg	9840
acaatggTgt	cccaatcctt	atacaaccaa	cttccacaag	aaagTcaagT	cagagacaa	9900
aaaaaaaaaa	gcaagggaaa	tttttttaatt	tgagTtTct	tgtttgctgc	ataatttatg	9960
cagTaaaaca	ctacacataa	cccttttagc	agtagagcaa	tggTtgaccg	tgtccttagc	10020

-continued

---

```

ttcttttatt ttattttttt atcagcaaag aataaataaa ataaatgag acacttcagg 10080
gatgtttcaa caagctctag agggccaat tcgacctata gtgagtcgta ttacaattca 10140
ctggccgtcg ttttacaacg tcgtgactgg gaaaaccctg gcgttacca acttaatcgc 10200
cttgacgac atcccccttt cgccagctgg cgtaatagcg aagaggccc caccgatcgc 10260
ccttccaac agttgcgag cctatac 10287

```

```

<210> SEQ ID NO 34
<211> LENGTH: 11480
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: vector pKR1149
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: (8496)..(8496)
<223> OTHER INFORMATION: n is a, c, g, or t

<400> SEQUENCE: 34

```

```

ggccgcaagt atgaactaaa atgcatgtag gtgtaagagc tcatggagag catggaatat 60
tgtatccgac catgtaacag tataataact gagctccatc tcacttcttc tatgaataaa 120
caaaggatgt tatgatatat taacctcta tctatgcacc ttattgttct atgataaatt 180
tcctcttatt attataaatc atctgaatcg tgacggctta tggaaatgctt caaatagtac 240
aaaaacaaat gtgtactata agactttcta aacaattcta accttagcat tgtgaacgag 300
acataagtgt taagaagaca taacaattat aatggaagaa gtttgcctcc atttatatat 360
tatatattac ccacttatgt attatattag gatgtaagg agacataaca attataaaga 420
gagaagtttg tatccattta tatattatat actaccatt tatatattat acttatccac 480
ttatttaatg tctttataag gtttgatcca tgatatttct aatattttag ttgatatgta 540
tatgaaaggg tactatttga actctcttac tctgtataaa ggttgatca tccttaaagt 600
gggtctattt aattttattg cttcttacag ataaaaaaaa aattatgagt tggtttgata 660
aaatattgaa ggatttaaaa taataataaa taacatataa tatatgtata taaatttatt 720
ataatataac atttatctat aaaaaagtaa atattgtcat aaatctatac aatcgtttag 780
ccttgctgga cgaatctcaa ttatttaaac gagagtaaac atatttgact ttttggttat 840
ttaacaaatt attatttaac actatatgaa attttttttt ttatcagcaa agaataaaat 900
taaattaaga aggacaatgg tgtccaatc cttatacaac caacttccac aagaaagtca 960
agtcagagac aacaaaaaaaa caagcaaagg aaatttttta atttgagttg tcttgtttgc 1020
tgcataatth atgcagtaaa acactacaca taaccttttt agcagtagag caatggttga 1080
ccgtgtgctt agcttctttt attttatttt tttatcagca aagaataaat aaaataaaat 1140
gagacacttc agggatgtht caacaagctc tagagggccc aattcgcctt atagtgagtc 1200
gtattacaat tcactggcgc tcgttttaca acgtcgtgac tgggaaaacc ctggcgcttac 1260
ccaacttaat cgccttgag cacatcccc ttctgccagc tggcgtaata gcgaagaggc 1320
ccgcaccgat cgccttctcc aacagttgag cagcctatac gtacgagatc cggccggcca 1380
gatcctgcag gagatccaag cttggcgcgc cgttctatag tgtcacctaa atcgtatgtg 1440
tatgatacat aaggttatgt attaattgta gcccggttct aacgacaata tgtccatattg 1500
gtgcactctc agtacaatct gctctgatgc cgcatagtta agccagcccc gacaccgccc 1560
aacaccgct gacgcgccct gacgggcttg tctgctcccc gcatccgctt acagacaagc 1620
gtgacccgtc tccgggagct gcatgtgta gaggttttca ccgtcatcac cgaacgcgc 1680

```

-continued

---

gagacgaaag	ggcctcgtga	tacgcctatt	tttatagggt	aatgtcatga	ccaaaatccc	1740
ttaacgtgag	ttttcgttcc	actgagcgtc	agaccccgta	gaaaagatca	aaggatcttc	1800
ttgagatcct	ttttttctgc	gcgtaactctg	ctgcttgcaa	acaaaaaac	caccgctacc	1860
agcggtggtt	tgtttgccgg	atcaagagct	accaactctt	tttccgaagg	taactggctt	1920
cagcagagcg	cagataccaa	atactgtcct	tctagtgtag	ccgtagttag	gccaccactt	1980
caagaactct	gtagcacccg	ctacatacct	cgctctgcta	atcctgttac	cagtggctgc	2040
tgccagtggc	gataagtcgt	gtcttaccgg	gttgactca	agacgatagt	taccggataa	2100
ggcgcagcgg	tcgggctgaa	cgggggggtc	gtgcacacag	cccagcttgg	agcgaacgac	2160
ctacaccgaa	ctgagatacc	tacagcgtga	gcattgagaa	agcgcaccgc	ttcccgaagg	2220
gagaaagcgg	gacaggtatc	cggtaagcgg	cagggtcggg	acaggagagc	gcacgagggg	2280
gcttccaggg	ggaacgcct	ggtatcttta	tagtcctgtc	gggtttcgcc	acctctgact	2340
tgagcgtcga	tttttgtgat	gctcgtcagg	ggggcggagc	ctatggaaaa	acgccagcaa	2400
cgcggccttt	ttacggttcc	tggccttttg	ctggcctttt	gctcacatgt	tctttcctgc	2460
gttatccct	gattctgtgg	ataaccgat	taccgccttt	gagtgagctg	ataccgctcg	2520
ccgcagccga	acgaccgagc	gcagcagagc	agtgagcgag	gaagcggaa	agcgcaccaat	2580
acgcaaaccg	cctctcccgc	cgcgttgccc	gattcattaa	tgcaggttga	tcgattcgac	2640
atcgatctag	taacatagat	gacaccgcgc	gcgataattt	atcctagttt	gcgcgctata	2700
ttttgttttc	tatcgcgtat	taaagtata	attgcgggac	tctaatacata	aaaaccctac	2760
tcataaataa	cgtcatgcat	tacatgtaa	ttattacatg	cttaacgtaa	ttcaacagaa	2820
attatatgat	aatcatcgca	agaccggcaa	caggattcaa	tcttaagaaa	ctttattgcc	2880
aatgtttga	acgatctgct	tcgacgcact	ccttctttag	gtacctcact	attcctttgc	2940
cctcggacga	gtgctggggc	gtcggtttcc	actatcggcg	agtacttcta	cacagccatc	3000
ggtccagacg	gccgcgcttc	tcggggcgat	ttgtgtacgc	ccgacagtcc	cggctccgga	3060
tcggacgatt	gcgtcgcctc	gaccctgcgc	ccaagctgca	tcctcgaaat	tgccgtcaac	3120
caagctctga	tagagttggt	caagaccaat	gcggagcata	tacgcccgga	gcccgggcga	3180
tcctgcaagc	tccggatgcc	tccgctcgaa	gtagcgcgtc	tgctgctcca	tacaagccaa	3240
ccacggcctc	cagaagaaga	tgttggcgac	ctcgtattgg	gaatcccga	acatcgctc	3300
gctccagtca	atgaccgctg	ttatgcggcc	attgtccgtc	aggacattgt	tggagccgaa	3360
atccgcgtgc	acgaggtgcc	ggacttcggg	gcagtcctcg	gcccaaagca	tcagctcatc	3420
gagagcctgc	gcgacggacg	cactgacggt	gtcgtccatc	acagtttgcc	agtgatacac	3480
atggggatca	gcaatcgcgc	atatgaaatc	acgccatgta	gtgtattgac	cgattccttg	3540
cggctccgaat	gggccgaacc	cgctcgtctg	gctaagatcg	gccgcagcga	tcgcatccat	3600
ggcctccgcg	accggctgca	gaacagcggg	cagttcggtt	tcaggcaggt	cttgcaacgt	3660
gacaccctgt	gcacggcggg	agatgcaata	ggtcaggctc	tcgctgaatt	ccccaatgtc	3720
aagcacttcc	ggaatcggga	gcgcggccga	tgcaaaagtgc	cgataaacat	aacgatcttt	3780
gtagaaacca	tcggcgcagc	tatttaccgg	caggacatat	ccacgcctc	ctacatcgaa	3840
gctgaaagca	cgagattctt	cgccctccga	gagctgcctc	aggtcggaga	cgctgtcgaa	3900
cttttcgatc	agaaacttct	cgacagacgt	cgcggtgagt	tcaggctttt	tcatggttta	3960
ataagaagag	aaaagagttc	ttttgttatg	gctgaagtaa	tagagaaatg	agctcgagcg	4020

-continued

---

tgtcctctcc	aatgaaatg	aacttcctta	tatagaggaa	gggtcctgcg	aaggatagtg	4080
ggattgtgcg	tcatocctta	cgtcagtgga	gatgtcacat	caatocactt	gctttgaaga	4140
cgtggttggg	acgtcttctt	tttccacgat	gctcctcgtg	gggtgggggc	catctttggg	4200
accactgtcg	gcagaggcat	cttgaatgat	agcctttcct	ttatcgcaat	gatggcattt	4260
gtaggagcca	ccttcctttt	ctactgtcct	ttcgatgaag	tgacagatag	ctgggcaatg	4320
gaatccgagg	aggtttcccg	aaattatcct	ttggtgaaa	gtctcaatag	ccctttggtc	4380
ttctgagact	gtatcttga	catttttggg	gtagaccaga	gtgtcgtgct	ccaccatgtt	4440
gacgaagatt	ttcttcttgt	cattgagtcg	taaaagactc	tgatgaact	gttcgccagt	4500
cttcacggcg	agttctgtta	gatcctcgat	ttgaatctta	gactccatgc	atggccttag	4560
attcagtagg	aactaccttt	ttagagactc	caatctctat	tacttgctt	ggtttatgaa	4620
gcaagccttg	aatcgccat	actggaatag	tacttctgat	cttgagaaat	atgtctttct	4680
ctgtgttctt	gatgcaatta	gtcctgaatc	ttttgactgc	atctttaacc	ttcttgggaa	4740
ggattttgat	ctcctggaga	ttgttactcg	ggtagatcgt	cttgatgaga	cctgctgctt	4800
aggcctctct	aaccatctgt	gggtcagcat	tctttctgaa	attgaagagg	ctaaccttct	4860
cattatcagt	gggtgaacata	gtgtcgtcac	cttcaccttc	gaacttcctt	cctagatcgt	4920
aaagatagag	gaaatcgtcc	attgtaatct	ccggggcaaa	ggagatctct	tttggggctg	4980
gatcactgct	gggccttttg	gttcctagcg	tgagccagtg	ggctttttgc	tttgggtggc	5040
ttgttagggc	cttagcaaa	ctcttgggct	tgagttgagc	ttctcctttg	gggatgaagt	5100
tcaacctgtc	tgtttctgta	cttgttgtgt	acgcgtcagc	tgctgctctt	gcctctgtaa	5160
tagtggcaaa	tttcttctgt	gcaactcccg	gaaacccgtt	tgttgccgcc	tttgtacaac	5220
cccagtcac	gtatataccg	gcattgtggac	cgttatacac	aacgtagtag	ttgatatgag	5280
gggtttgaat	acccgattct	gctctgagag	gagcaactgt	gctgttaagc	tcagattttt	5340
gtgggattgg	aattggatcg	atctcgatcc	cgcgaaatta	atacgactca	ctatagggag	5400
accacaacgg	tttccctcta	gaaataattt	tgtttaactt	taagaaggag	atatacccat	5460
ggaaaagcct	gaactcaccg	cgacgtctgt	cgagaagttt	ctgatcgaaa	agttcgacag	5520
cgtctccgac	ctgatgcagc	tctcggaggg	cgaagaatct	cgtgctttca	gcttcgatgt	5580
aggagggcgt	ggatatgtcc	tgccggtaaa	tagctgcgcc	gatggtttct	acaaagatcg	5640
ttatgtttat	cggcactttg	catcggccgc	gctcccgatt	ccggaagtgc	ttgacattgg	5700
ggaattcagc	gagagcctga	cctattgcat	ctcccgcctg	gcacaggggtg	tcacgttgca	5760
agacctgcct	gaaaccgaac	tgcccctgtg	tctgcagccg	gtcgcggagg	ctatggatgc	5820
gatcgcctcg	gccgatctta	gccagacgag	cgggttcggc	ccattcggac	cgcaaggaa	5880
cggccaatac	actacatggc	gtgatttcat	atgcgcgatt	gctgatcccc	atgtgtatca	5940
ctggcaaaact	gtgatggacg	acaccgtcag	tgcgctcgtc	gcgcaggctc	tcgatgagct	6000
gatgctttgg	gccgaggact	gccccgaagt	ccggcacctc	gtgcacgcgg	atttcggctc	6060
caacaatgtc	ctgacggaca	atggccgcat	aacagcggtc	attgactgga	gcgaggcgat	6120
gttcggggat	tccaataacg	aggctcgcaa	catcttcttc	tggaggccgt	ggttggcttg	6180
tatggagcag	cagacgcgct	acttcgagcg	gaggcatccg	gagcttgacg	gatcgcgcg	6240
gctccggggc	tatatgtctc	gcattggctt	tgaccaactc	tatcagagct	tggttgacgg	6300
caatttcgat	gatgcagctt	gggcgcaggg	tcgatgcgac	gcaatcgtcc	gatccggagc	6360
cgggactgtc	gggcgtacac	aaatcgcccg	cagaagcgcg	gccgtctgga	ccgatggctg	6420

-continued

---

tgtagaagta	ctcgccgata	gtggaaaccg	acgccccagc	actcgtccga	gggcaaagga	6480
atagtgaggt	acagcttga	tcgatccggc	tgctaacaaa	gccccaaagg	aagctgagtt	6540
ggctgctgcc	accgctgagc	aataactagc	ataaccctt	ggggcctcta	aacgggtcct	6600
gaggggtttt	ttgctgaaag	gaggaactat	atccggatga	tcgggcgcgc	cgctcgacga	6660
tccactagtt	ctagagcggc	ccgcgcgctc	gacggatata	atgagccgta	aacaaagatg	6720
attaagtagt	aattaatcag	tactagtaaa	agtggcaaaa	gataacgaga	aagaaccaat	6780
ttctttgcat	tcggccttag	cgaaggcat	atataagctt	tgattatfff	atftagtgt	6840
atgatttcgt	acaaccaaag	catttattta	gtactctcac	acttgtgtcg	cggccggccg	6900
ctacaggaac	agggtggtgg	ggccctcggc	gcgctcgtac	tgctccacga	tggtgtagtc	6960
ctcgtgtggt	gaggtgatgt	ccagcttga	gtccacgtag	tagtagccgg	gcagctgcac	7020
gggcttcttg	gccatgtaga	tggactgaa	ctccaccagg	tagtggccgc	cgctcctcag	7080
cttcagggcc	ttgtggatct	cgccctcag	cacgcgctcg	cggggtaca	ggcgctcgg	7140
ggaggcctcc	cagcccatag	tcttctctg	cattacgggg	ccgtcggagg	ggaagttcac	7200
gccgatgaac	ttcaccttgt	agatgaagga	gccgtcctgc	agggaggagt	cctgggtcac	7260
ggtcaccacg	ccgcccctc	cgaagtcat	cacgcgctcc	cacttgaagc	cctcggggaa	7320
ggacagcttc	ttgtagtccg	ggatgtcggc	gggtgtcttc	acgtacacct	tgagcgcgta	7380
ctggaactgg	ggggacagga	tgtcccaggc	gaagggcagg	gggcccctc	tggtcacctt	7440
cagcttgccg	gtctgggtgc	cctcgtagg	gcggccctcg	ccctcgcctc	cgatctcgaa	7500
ctcgtggccg	ttcacggagc	cctccatgcg	caccttgaag	cgcatgaact	ccttgatgac	7560
gtcctcggag	gaggccatgg	gccgctggg	gggctatgga	agactttctt	agttagtgt	7620
gtgaataaag	aatgttggga	gaatcgggac	tacttatagg	ataggaataa	aacagaaaag	7680
tattaagtgc	taatgaaata	tttagactga	taattaaaat	cttcacgtat	gtccacttga	7740
tataaaaacg	tcaggaataa	aggaagtaca	gtagaattta	aaggtaactc	ttttatata	7800
accctgttcc	tctttttggc	tagctagtgg	cataaaaaat	aatctatatt	tttatcatta	7860
ttttaaat	cttatgagat	ggtaaatatt	tatcataatt	tttttacta	ttatttatta	7920
tttgtgtgtg	taatacatat	agaagttaat	tacaaatfff	atttactfff	tcattatfff	7980
gatatgatcc	accattaatt	tagtgttatt	atftataata	gttcatttta	atctttttgt	8040
atataattatg	cggtgcagtac	tttttctc	catataacta	ctattacatt	ttatttata	8100
aatatftt	ttaatgaatt	ttcgtgataa	tatgtaatat	tgttcattat	tatttcagat	8160
tttttaaaaa	tatttgtgtt	attatfttag	aaatatgtaa	tttttttagt	atttgatftt	8220
atgatgataa	agtgttctaa	attcaaaaga	agggggaaag	cgtaaacatt	aaaaaacgtc	8280
atcaaaaa	aacaaaatct	tgtaataaa	gataaaactg	tttgttttga	tcactgttat	8340
ttcgtaatat	aaaaacatta	tttatattta	tattgttgac	aaccaaattt	gcctatcaaa	8400
tctaaccaat	ataatgcag	cgtggcaggt	aatgtactac	catgaactta	agtcagaca	8460
taataaacgg	tgaatctgac	caatgcaggt	acctanctaa	attgtatttg	tgacacgaag	8520
caaatgatcc	aattcacaat	ggagatggga	aacaaataat	gaagaacca	gaactaagaa	8580
agcttttctg	aaaaataaaa	taaaggcaat	gtcaaaagta	tactgcatca	tcagtccaga	8640
aagcacatga	tattttttta	tcagtatcaa	tcagctaggt	tttattttac	aatatcgata	8700
tagctagttt	aaatatattg	cagctagatt	tataaatatt	tgtgttatta	tttatcattt	8760

-continued

---

gtgtaatcct	gttttttagta	tttttagtta	tatatgatga	taatgtattc	caaatttaaa	8820
agaagggaaa	taaattttaaa	caagaaaaaa	agtcatcaaa	caaaaaacaa	atgaaaggggt	8880
ggaaagatgt	taccatgtaa	tgtgaatggt	acagtatttc	ttttattata	gagttaacaa	8940
attaactaat	atgattttgt	taataatgat	aaaatatttt	ttttattatt	atttcataat	9000
ataaaaatag	tttacttaat	ataaaaaaaa	ttctatcggt	cacaacaaag	ttggccacct	9060
aatttaacca	tgcattgtacc	catggaccat	attaggtaac	catcaaacct	gatgaagaga	9120
taaagagatg	aagacttaag	tcataacaca	aaaccataaa	aaacaaaaat	acaatcaacc	9180
gtcaatctga	ccaatgcatg	aaaaagctgc	aatagtgagt	ggcgacacaa	agcacatgat	9240
tttcttacaa	cggagataaa	acaaaaaaa	tatttctatg	acaacctaga	acaaataaag	9300
cttttatata	ataaatatat	aaataaataa	aggctatgga	ataatatact	tcaatatatt	9360
tggattaaat	aaattgttgg	cggggttgat	atattttatac	acacctaaag	tcacttcaat	9420
ctcattttca	cttaactttt	attttttttt	tctttttatt	tatcataaag	agaatattga	9480
taataacttt	tttaacatat	ttttatgaca	ttttttattg	gtgaaaactt	attaaaaatc	9540
ataaattttg	taagttgat	ttattttaaag	agttcctcct	cttattttta	attttttaat	9600
aaatttttaa	ataactaaaa	tttgtgttaa	aaatgttaa	aaagtgtgtt	attaaccctt	9660
ctcttcgagg	atccgtaccg	agctcggatc	ctctagaaat	ccgtcaacat	ggtggagcac	9720
gacactctcg	tctactccaa	gaatatcaaa	gatacagtct	cagaagacca	aagggtctatt	9780
gagacttttc	aacaaaggggt	aatatcggga	aacctcctcg	gattccattg	cccagctatc	9840
tgtcacttca	tcaaaaggac	agtagaaaag	gaaggtggca	cctacaaatg	ccatcattgc	9900
gataaaggaa	aggctatcgt	tcaagatgcc	tctgcgcaca	gtggtcccaa	agatggaccc	9960
ccaccaccga	ggagcatcgt	ggaaaaagaa	gacgttccaa	ccacgtcttc	aaagcaagtg	10020
gattgatgtg	atgatcctat	gcgtatggta	tgactgtgtg	tcaagatgat	gacttcaaac	10080
ctacctatga	cgtatggtat	gaactgtgtg	cgactgatga	cttagatcca	ctcgagcggc	10140
tataaatacg	tacctacgca	cctgcgcgta	ccatccctag	agctgcagct	tattttttaca	10200
acaattacca	acaacaacaa	acaacaaaca	acattacaat	tactatttac	aattacagtc	10260
gaccgggat	cgtacctcta	gggtggcggc	cgcattggaga	gatctcaacg	gcagctcctc	10320
ccgccaccgt	cgcgcgtctc	ctcctcgtcc	tccgtctccg	cggacaccgt	cctcgtcctc	10380
cccggaaaga	ggcggagggc	ggcgcagggc	aaggccggcg	ccgagcctaa	taagaggatc	10440
cgcaaggacc	ccgccgcgcg	cgccgcgggg	aagaggagct	ccgtctacag	gggagtcacc	10500
aggcacaggt	ggacgggag	ggtcgaggcg	catctctggg	acaagcactg	cctcgccgcg	10560
ctccacaaca	agaagaaagg	caggcaagtc	tacctggggg	cgtatgacag	cgaggaggca	10620
gctgctcgtg	cctatgacct	cgcagctctc	aagtactggg	gtcctgagac	tctgctcaac	10680
ttcctcgtgg	aggattactc	cagcagatg	ccggagatgg	aggccgtgct	ccgggaggag	10740
tacctggcct	ccctccgcgc	caggagcagc	ggcttctcca	ggggcgtctc	caagtacaga	10800
ggcgtcgcca	ggcatcacca	caacgggagg	tgggaggcgc	ggattggggc	agtctttggg	10860
aacaagtacc	tctacttggg	aacatttgac	actcaagaag	aggcagccaa	ggcctatgac	10920
cttgccggcca	ttgaataccg	tggcgtcaat	gctgtaacca	acttcgacat	cagctgctac	10980
ctggaccacc	cgtgttctct	ggcacagctc	caacaggagc	cacaggtggg	gccggcactc	11040
aaccaagaac	ctcaacctga	tcagagcgaa	accggaacta	cagagcaaga	gccggagtca	11100
agcgaagcca	agacaccgga	tggcagtgca	gaaccgatg	agaacgcggg	gcctgacgac	11160

-continued

---

```

accgcgggagc ccctcaccac agtcgaagac agcatcgaag agggccttggtg gagcccttgc 11220
atggattacg agctagacac catgctcaga ccaaactttg gcagctcaat caatctgagc 11280
gagtggttcg ctgacgcaga cttcgactgc aacatcggat gcctgttcga tgggtgttct 11340
gcggtcgaag aaggaagcaa ggtggtgta ggtctggcag atttcagtct gtttgaggca 11400
ggtgatgtcc agctgaagga tgtctcttcg gatatggaag aggggatata acctccagcg 11460
atgatcagtg tgtgcaacgc 11480

```

```

<210> SEQ ID NO 35
<211> LENGTH: 19472
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: vector pKR1221
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: (15083)..(15083)
<223> OTHER INFORMATION: n i s a , c , g , o r t

```

```

<400> SEQUENCE: 35

```

```

cgcgccagat cctctagagt cgacctgcag gcatgcaagc ttggcgtaat catggtcata 60
gctgtttcct gtgtgaaatt gttatcgcct cacaattcca cacaacatac gagccggaag 120
cataaagtgt aaagcctggg gtgcctaatt agtgagctaa ctcacattaa ttgctgtgag 180
ctcactgccc gctttccagt cgggaaacct gtcgtgccag ctgcattaat gaatcggcca 240
acgcgcgggg agaggcgggt tgcgtattgg atcgatccct gaaagcgacg ttggatgta 300
acatctacaa attgcctttt cttatcgacc atgtacgtaa gcgcttacgt ttttggtgga 360
cccttgagga aactggtagc tgttgtgggc ctgtggtctc aagatggatc attaatttcc 420
accttcacct acgatggggg gcatcgacc ggtgagtaat attgtacggc taagagcgaa 480
tttgccctgt agacctcaat tgcgagcttt ctaatttcaa actattcggg cctaactttt 540
ggtgtgatga tgctgactgg caggatatac accgttgtaa tttgagctcg tgtgaataag 600
tcgctgtgta tgtttgtttg attgtttctg ttggagtgca gccatttca cgggacaagt 660
cggetagatt gatttagccc tgatgaactg ccgaggggaa gccatcttga gcgcggaatg 720
ggaatggatt tcgttgtaca acgagacgac agaaccacca cgggaccgag cttcgcgagc 780
ttttgatccc gtgcatcctc tggccgggc gatttgttca cgtccatgag gcgctctcca 840
aaggaaacga tttttccggg tgaaccttt ccggttcttc ctctactcga cctcttgaag 900
tcccagcatg aatgttcgac cgtcccgcaa gcgcatcttt ggcgcaacca gccggtttcg 960
cacgtcgatt ctcgcgagcc tgcatacttt ggcaagattg ctgaatgacg ctgatgcttc 1020
atcgcaatct gcgataatgg ggtaaagtac cgggtgaaggc cgcaggtcag gccgcctgag 1080
cactcagtgt cttggatgtc cagttccacg gcagctgttg ctcaagcctg ctgatcggag 1140
cgtccgcaag gtcggcgcgg acgtcggcaa gccaggcctg cggatcagtg ttattgagct 1200
tggcgctcat gatcagtgtc gccatgaacg ccgacgcttc agcacaacga tccgatccgg 1260
caaacagcca tgacttctcg ccgagtacat agcctctgag cgttcgttcc gcagcattgt 1320
tcgtcaggca aatcgggccc tcactcagga atgacgtaat gccatcccat cgcttgagca 1380
tgtaatttat gcctcggcg acgggagaaac tgccgcgaaa tttccccgc tcggtttcga 1440
gccaatcatg cagctcttcg gcgagtgacc ttgatcaggc caccgccacg accgcggaag 1500
acgaacagat gcctgcgcat cggatcgcgc ttcagcgtct cttgcaacct cagcgacaaa 1560

```

-continued

---

ccgggaaagc ctttgcgcac gtcctgactt atgtcgcac ttgggggggc ttcgtctacg	1620
tggccttcgt gategacgtc ttcgccctgc gcattgtcgg atggcgggcg agccggacag	1680
cacatgcagg ctttgcctc gatgccctcg aggaggctca tcatgatcgg cgtcccgcctc	1740
atggcggcct agtgcacac tcggatcgcg gtgttcaata cgtgtccttt cgctattccg	1800
agcggttggc agaagcaggt atcgagccat ctatcggaag cgtcggcgac agcacgacaa	1860
cgccctcgca gaagcgatca acggctctta caaggccgag gtcattcacc ggcgtggacc	1920
atggaggagc ttcgaagcgg tcgagttcgc taccttggaa tggatagact ggttcaacca	1980
cggcggtttt tgaagcccat cggcaatata ccgccagccg aagacgagga tcagtattac	2040
gccatgctgg acgaagcagc catggctgcg cattttaacg aaatggcctc cggcaaaccc	2100
ggtgcggttc acttgttgcg tgggaaagt caccggactc cgcgcacgag ccttctctgt	2160
aatagccata tcgaccgaat tgacctgcag gggggggggg gaaagccacg ttgtgtctca	2220
aaatctctga tgttacattg cacaagataa aaatatatca tcatgaacaa taaaactgtc	2280
tgcttacata aacagtaata caaggggtgt tatgagccat attcaacggg aaacgtcttg	2340
ctcagggcgg cgattaaatt ccaacatgga tgctgattta tatgggtata aatgggctcg	2400
cgataaatgtc gggcaatcag gtgcgacaat ctatcgattg tatgggaagc ccgatgcgcc	2460
agagttgttt ctgaaacatg gcaaaggtag cgttgccaat gatgttacag atgagatggt	2520
cagactaac tcggctgacgg aatttatgcc tcttcgacc atcaagcatt ttatccgtac	2580
tctgatgat gcattggtac tcaccactgc gatccccggg aaaacagcat tccaggtatt	2640
agaagaatat cctgattcag gtgaaaatat tgttgatgcg ctggcagtg tcttgcgccg	2700
gttgattcgg attcctgttt gtaattgtcc ttttaacagc gatcgcgtat ttcgtctcgc	2760
tcagggcgaa tcacgaatga ataacggttt ggttgatgcg agtgattttg atgacgagcg	2820
taatggctgg cctgttgaac aagtctgga agaaatgcat aagcttttgc cattctcacc	2880
ggattcagtc gtcactcatg gtgattttc acttgataac cttatttttg acgaggggaa	2940
attaataggt tgtattgatg ttggacgagt cggaaatcga gaccgatacc aggatcttgc	3000
catcctatgg aactgcctcg gtgagttttc tcttcatta cagaacggc tttttcaaaa	3060
atattggtatt gataatcctg atatgaataa attgcagttt catttgatgc tcgatgagtt	3120
tttctaatca gaattggtta attggttgta acaactggcag agcattacgc tgacttgacg	3180
ggacggcggc tttgttgaat aaatcgaact tttgctgagt tgaaggatca gatcagcat	3240
cttcccagca acgcagaccg ttcctgtggca aagcaaaagt tcaaaatcac caactggtcc	3300
acctacaaca aagctctcat caaccgtggc tccctcactt tctggctgga tgatggggcg	3360
attcagccct ggtatgagtc agcaaacct tcttcacgag gcagacctca gcgcccccc	3420
ccccctcag gtcttttcca atgatgagca cttttaaagt tctgctatgt ggcgcgggat	3480
tatcccgtgt tgacgccggg caagagcaac tcggctgccc catacactat tctcagaatg	3540
acttggttga gtactcacca gtcacagaaa agcatcttac ggatggcatg acagtaagag	3600
aattatgcag tgctgccata accatgagtg ataacactgc ggccaactta cttctgacaa	3660
cgatcggagg accgaaggag ctaaccgctt ttttgacaaa catgggggat catgtaactc	3720
gccttgatcg ttgggaaccg gagctgaatg aagccatacc aaacgacgag cgtgacacca	3780
cgatgcctgt agcaatggca acaacgttgc gcaaacatt aactggcgaa ctacttactc	3840
tagcttcccg gcaacaatta atagactgga tggaggcggg taaagttgca ggaccacttc	3900
tgcgctcggc ccttccggct ggtcgttcta ttgctgataa atctggagcc ggtgagcgtg	3960

-continued

---

ggtctcgcgg	tatcattgca	gcaactgggc	cagatggtaa	gccctcccg	atcgtagtta	4020
tctacacgac	ggggagtcag	gcaactatgg	atgaacgaaa	tagacagatc	gctgagatag	4080
gtgcctcact	gattaagcat	tggttaactgt	cagaccaagt	ttactcatat	atactttaga	4140
ttgatttaaa	acttcatttt	taatttaaaa	ggatctaggt	gaagatcctt	tttgataatc	4200
tcatgaccaa	aatcccttaa	cgtgagtttt	cgttccactg	agcgtcagac	cccgtagaaa	4260
agatcaaagg	atcttcttga	gatocttttt	ttctgcgcgt	aatctgctgc	ttgcaaaaa	4320
aaaaaccacc	gctaccagcg	gtggtttgtt	tgccggatca	agagctacca	actctttttc	4380
cgaaggtaac	tggcttcagc	agagcgcaga	taccaaatac	tgtccttcta	gtgtagccgt	4440
agttaggcca	ccacttcaag	aactctgtag	caccgcctac	atacctcgct	ctgctaatac	4500
tgttaccagt	ggctgctgcc	agtggcgata	agtcgtgtct	taccgggttg	gactcaagac	4560
gatagttaac	ggataaggcg	cagcggtcgg	gctgaacggg	gggttcgtgc	acacagccca	4620
gcttggagcg	aacgacctac	accgaactga	gatacctaca	gcgtgagcta	tgagaaagcg	4680
ccacgcttcc	cgaagggaga	aagggcgaca	ggtatccggt	aagcggcagg	gtcggaaacg	4740
gagagcgcac	gagggagctt	ccaggggaa	acgcctggta	tctttatagt	cctgctgggt	4800
ttcgccacct	ctgacttgag	cgctgatttt	tgtgatgctc	gtcagggggg	cggagcctat	4860
ggaaaaacgc	cagcaacgcg	gocctttttac	ggttcctggc	cttttgetgg	ccttttgctc	4920
acatgttctt	tctctcgctta	tcccctgatt	ctgtggataa	ccgtattacc	gcctttgagt	4980
gagctgatac	cgctcgcgcg	agccgaacga	ccgagcgcag	cgagtcagtg	agcagggaa	5040
cggaaagagcg	cctgatgocg	tattttctcc	ttacgcctct	gtgocgtatt	tcacaccgca	5100
tatggtgcac	tctcagtaca	atctgctctg	atgccgcata	gttaagccag	tatacactcc	5160
gctatcgcta	cgtgactggg	tcattgctgc	gccccgacac	ccgccaacac	ccgctgacgc	5220
gccctgacgg	gcttgtctgc	tcccggcctc	cgcttacaga	caagctgtga	ccgtctccgg	5280
gagctgcatg	tgtcagaggt	tttcaccgtc	atcacgaaaa	cgcgcgaggg	agggtgccct	5340
gatgtgggcg	ccggcggctg	agtggcgacg	gcgcggcctg	tcgcgcacct	ggtagattgc	5400
ctggccgtag	gccagccatt	tttgagcggc	cagcggcccg	gataggccga	cgcgaagcgg	5460
cggggcgtag	ggagcgcagc	gaccgaaggg	taggcgcttt	ttgcagctct	tcggctgtgc	5520
gctggccaga	cagttatgca	caggccaggc	gggttttaag	agttttaata	agttttaaag	5580
agttttaggg	ggaaaaatcg	ccttttttct	cttttatatc	agtcacttac	atgtgtgacc	5640
ggttcccaat	gtacggcttt	gggttcccaa	tgtacgggtt	ccggttccca	atgtacggct	5700
ttgggttccc	aatgtacgtg	ctatccacag	gaaagagacc	ttttcgacct	ttttcccctg	5760
ctagggcaat	ttgcctagc	atctgctccg	tacattagga	accggcggat	gcttgcacct	5820
cgatcaggtt	gcggtagcgc	atgactagga	tcgggccagc	ctgccccgcc	tcctccttca	5880
aatcgtactc	cggcaggtca	tttgaccgca	tcagcttgcg	cacggtgaaa	cagaacttct	5940
tgaactctcc	ggcgctgcca	ctgcgttctg	agategtctt	gaacaacat	ctggcttctg	6000
ccttgctctc	ggcgccggct	gccaggcggg	agagaaaaag	gccgatgccg	ggatcgatca	6060
aaaagtaatc	ggggtaaac	gtcagcacgt	ccgggttctt	gccttctgtg	atctcgcggg	6120
acatccaatc	agctagctcg	atctcgatgt	actccggccg	cccggttctg	ctctttacga	6180
tctttagcgc	gctaatcaag	gcttaccctc	cggataccgt	caccagcggg	ccgttcttgg	6240
ccttctctct	acgctgcatg	gcaacgtgcg	tgggttttaa	ccgaatgcag	gtttctacca	6300

-continued

---

ggtcgtcttt	ctgctttccg	ccatcggttc	gccggcagaa	cttgagtacg	tccgcaacgt	6360
gtggacggaa	cacgcggccg	ggcttgtctc	ccttcccttc	ccggtatcgg	tccatggatt	6420
cggtagatg	ggaaaccgcc	atcagtacca	ggtcgtaatc	ccacacactg	gccatgccgg	6480
ccggccctgc	ggaaacctct	acgtgccctg	ctggaagctc	gtagcggatc	acctcgccag	6540
ctcgtcggtc	acgcttcgac	agacggaaaa	cggccaacgtc	catgatgctg	cgactatcgc	6600
gggtgcccac	gtcatagagc	atcggaacga	aaaaatctgg	ttgctcgtcg	cccttggggc	6660
gcttcctaata	cgacggcgca	ccggctgccc	gcggttgccc	ggattctttg	cggattcgat	6720
cagcggccgc	ttgccacgat	tcaccggggc	gtgcttctgc	ctcgtatcgt	tgcgctggg	6780
cggctcgcgc	ggccttcaac	ttctccacca	ggtcacacc	cagcggccgc	cggatttgta	6840
ccgggcccga	tggtttgoga	ccgctcacgc	cgattcctcg	ggcttggggg	ttccagtgcc	6900
attgcagggc	cggcagacaa	cccagccgct	tacgcctggc	caaccgccc	ttctccaca	6960
catggggcat	tccacggcgt	cggtgccctg	ttgttcttga	ttttccatgc	cgctccttt	7020
agccgctaaa	attcatctac	tcatttattc	atttgcctat	ttactctggg	agctgcgcga	7080
tgtattcaga	tagcagctcg	gtaatggtct	tgccttggcg	taccgcgtac	atcttcagct	7140
tggtgtgac	ctccgcccgc	aactgaaagt	tgaccgctt	catggctggc	gtgtctgcca	7200
ggctggccaa	cgttgcagcc	ttgctgctgc	gtgcgctcgg	acggcccggc	cttagcgtgt	7260
ttgtgctttt	gctcattttc	tctttacctc	attaactcaa	atgagttttg	atttaatttc	7320
agcggcccagc	gctcggacct	cgcgggcagc	gtcgcctcgc	ggttctgatt	caagaacggt	7380
tgtgccggcg	gcggcagctg	ctgggtagct	cacgcgctgc	gtgatacggg	actcaagaat	7440
gggcagctcg	taccggccca	gcgcctcggc	aacctcaccc	ccgatgcgcg	tgcctttgat	7500
cgcccgcgac	acgacaaagg	ccgctttag	ccttccatcc	gtgacctcaa	tgcgctgctt	7560
aaccagctcc	accaggtcgg	cgggtggcca	tatgtcgtaa	gggcttggct	gcaccggaat	7620
cagcacgaag	tcggctgcct	tgatcgcgga	cacagccaag	tcgcgcgcct	ggggcgctcc	7680
gtcgatecact	acgaagtgcg	gccggccgat	ggccttcacg	tcgcgggtcaa	tcgtcggggc	7740
gtcgatecgc	acaacggtta	gcggttgatc	ttcccgcacg	gccgcccact	cgcgggcaact	7800
gccctgggga	tcggaatcga	ctaacagAAC	atcggccccg	gcgagttgca	gggcgcgggc	7860
tagatggggt	gcgatggtcg	tcttgcttga	cccgcctttc	tggttaagta	cagcgataac	7920
ttcatgcgtt	cccttgcgta	ttgtttatt	tactcatcgc	atcatatcgc	cagcgaccgc	7980
atgacgcaag	ctgttttact	caaatacaca	tcacctttt	agacggcggc	gctcggtttc	8040
ttcagcggcc	aagctggccg	gccaggccgc	cagcttggca	tcagacaaac	cggccaggat	8100
ttcatgcagc	cgcacggttg	agacgtgcgc	gggcggctcg	aacacgtacc	cggcccgcat	8160
catctccgcc	tcgatctcct	cggtaatgaa	aaacggttcg	tcctggccgt	cctgggtcgg	8220
tttcatgctt	gttctcttg	gcgttcattc	tcggcggccg	ccagggcgtc	ggcctcggtc	8280
aatgcgtcct	cacggaagge	accgcgccgc	ctggcctcgg	tgggcgtcac	ttcctcgtcg	8340
cgctcaagtg	cgcggtacag	ggtcgagcga	tgcacgcaaa	gcagtgcagc	cgctctttc	8400
acggtgcggc	cttctcggtc	gatcagctcg	cgggcgtgcg	cgatctgtgc	cggggtgagg	8460
gtagggcggg	ggccaaactt	cacgcctcgg	gccttggcgg	cctcgcgccc	gctccgggtg	8520
cggctcgatg	ttagggaacg	ctcgaactcg	gcaatgccgg	cgaacacggt	caacaccatg	8580
cggccggccg	gcgtggtggt	gtcggccacc	ggctctgcca	ggctacgcag	gcccgcccg	8640
gcctcctgga	tgcgctcggc	aatgtccagt	aggtcgcggg	tgctgcgggc	caggcgtct	8700

-continued

---

agcctggta	ctgtcacaac	gtcgccaggg	cgtaggtggt	caagcatcct	ggccagctcc	8760
gggcggtcgc	gcctgggtgcc	ggtgatcttc	tcggaaaaca	gcttgggtgca	gccggccgcg	8820
tgcaagtccg	cccgttggtt	ggtcaagtcc	tggtcgctcg	tgctgacgcg	ggcatagccc	8880
agcaggccag	eggcggcgct	cttgttcatg	gcgtaatgic	tccggttcta	gtcgcaagta	8940
ttctacttta	tgcgactaaa	acacgcgaca	agaaaacgcc	aggaaaaggg	cagggcggca	9000
gcctgtcgcg	taacttagga	cttggtcgac	atgtcgtttt	cagaagacgg	ctgcaactgaa	9060
cgtcagaagc	cgactgcact	atagcagcgg	aggggttgga	ccacaggacg	ggtgtggtcg	9120
ccatgatcgc	gtagtcgata	gtggctccaa	gtagcgaagc	gagcaggact	gggcggcggc	9180
caaagcggtc	ggacagtgtc	ccgagaacgg	gtgcgcatag	aaattgcatc	aacgcatata	9240
gcgctagcag	cacgccatag	tgactggcga	tgctgtcgga	atggacgata	tcccgaaga	9300
ggcccggcag	taccggcata	accaagccta	tgctacagc	atccagggtg	acggtgccga	9360
ggatgacgat	gagcgcattg	ttagatttca	tacacgggtc	ctgactgcgt	tagcaattta	9420
actgtgataa	actaccgcat	taaagctagc	ttgcttggtc	gttccgcgtg	aacgtcggtc	9480
cgattgtaac	tgcgttcaaa	tactttgca	tcgtgttcg	cgctgcccg	gtcgctcggc	9540
tgatctcacg	gatcgactgc	ttctctcgca	acgccatccg	acggatgatg	tttaaaagtc	9600
ccatgtggat	cactccgttg	ccccgtcgt	caccgtgttg	gggggaaggt	gcacatggct	9660
cagttctcaa	tggaaattat	ctgcctaacc	ggctcagttc	tgctagaaa	ccaacatgca	9720
agctccacgc	ggtgcaaagc	ggcagcggcg	gcaggatata	ttcaattgta	aatggcttca	9780
tgccgggaa	atctacatgg	atcagcaatg	agtatgatgg	tcaatatgga	gaaaaagaaa	9840
gagtaattac	caattttttt	tcaattcaaa	aatgtagatg	tccgcagcgt	tattataaaa	9900
tgaaagtaca	ttttgataaa	acgacaaatt	acgatccgtc	gtatttatag	gcgaaagcaa	9960
taaacaaatt	attctaattc	ggaaatcttt	atctcgactg	gtctacattc	acgtccaaat	10020
gggggcttag	atgagaaact	tcacgatcga	tgcttgatt	tcgcoattcc	cagataccca	10080
tttcatcttc	agattggtct	gagattatgc	gaaaatatac	actcatatac	ataaatactg	10140
acagtttgag	ctaccaattc	agtgtagccc	attacctcac	ataattcact	caaatgctag	10200
gcagtctgtc	aactcggcgt	caatttgctg	gccactatac	gatagttgctg	caaattttca	10260
aagtcctggc	ctaacatcac	acctctgtcg	gcggcgggtc	ccatttgtga	taaatccacc	10320
atatcgaatt	aattcagact	cctttgcccc	agagatcaca	atggacgact	tcctctatct	10380
ctacgatcta	gtcaggaagt	tcgacggaga	aggtgacgat	accatgttca	ccactgataa	10440
tgagaagatt	agccttttca	atctcagaaa	gaatgctaac	ccacagatgg	ttagagaggc	10500
ttacgcagca	ggtctcatca	agacgatcta	cccagcaat	aatctccagg	agatcaataa	10560
ccttccaag	aaggttaaag	atgcagtcaa	aagattcagg	actaactgca	tcaagaacac	10620
agagaagat	atatttttca	agatcagaag	tactattcca	gtatggacga	ttcaaggctt	10680
gcttcacaaa	ccaaggcaag	taatagagat	tggagtctct	aaaaaggtag	ttccactga	10740
atcaaaggcc	atggagtcaa	agattcaaat	agaggaccta	acagaactcg	ccgtaagac	10800
tggcgaacag	ttcatacaga	gtctcttacg	actcaatgac	aagaagaaaa	tcttcgtaaa	10860
catggtggag	cacgacacgc	ttgtctactc	caaaaatatac	aaagatacag	tctcagaaga	10920
ccaaaggcca	attgagactt	ttcaacaaag	ggtaatatcc	ggaacctcc	tcggattcca	10980
ttgccagct	atctgtcact	ttattgtgaa	gatagtgaa	aaggaaggtg	gctcctacaa	11040

-continued

---

atgccatcat	tgcgataaag	gaaaggccat	cgttgaagat	gcctctgccg	acagtgggcc	11100
caaagatgga	ccccaccoca	cgaggagcat	cgtggaaaaa	gaagacgttc	caaccacgtc	11160
ttcaaagcaa	gtggattgat	gtgatatctc	cactgacgta	agggatgacg	cacaatoccca	11220
ctatcctctg	caagaccott	cctctatata	aggaagtcca	tttcatttgg	agaggacacg	11280
ctgaaatcac	cagtctccaa	gcttgccggg	atcgtttcgc	atgattgaac	aagatggatt	11340
gcacgcaggt	tctccggcgc	cttgggtgga	gaggetatct	ggctatgact	gggcacaaca	11400
gacaatcggc	tgctctgatg	ccgccgtgtt	ccggctgtca	gcgcaggggc	gcccggttct	11460
ttttgtcaag	accgacatgt	ccggtgccct	gaatgaactg	caggacgagg	cagcgcggct	11520
atcgtggctg	gccacgacgg	gcgttccttg	cgcagctgtg	ctcgacgttg	tactgaagc	11580
gggaagggac	tggctgctat	tgggcgaagt	gccggggcag	gatctcctgt	catctcacct	11640
tgctcctgcc	gagaaagtat	ccatcatggc	tgatgcaatg	cgccggctgc	atacgttga	11700
tccggctacc	tgcccattcg	accaccaagc	gaaacatcgc	atcgagcggg	cacgtactcg	11760
gatggaagcc	ggtcttctgc	atcaggatga	tctggacgaa	gagcatcagg	ggctcgcgcc	11820
agccgaactg	tccgccaggc	tcaaggcgcg	catgcccgac	ggcgaggatc	tcgtcgtgac	11880
ccatggcgat	gcctgcttgc	cgaatatcat	ggtggaaaaa	ggccgctttt	ctggattcat	11940
cgactgtggc	cggctgggtg	tggcggaccg	ctatcaggac	atagcgttgg	ctaccctgga	12000
tattgctgaa	gagcttggcg	gcgaatgggc	tgaccgcttc	ctcgtgcttt	acggatcgc	12060
cgctcccgat	tcgcagcgca	tcgccttcta	tcgccttctt	gacgagttct	tctgagcggg	12120
actctggggg	tcgaaatgac	cgaccaagcg	acgcccacc	tgccatcacg	agatttcgat	12180
tccaccgcgc	ccttctatga	aaggttgggc	ttcggaatcg	ttttccggga	cgccggctgg	12240
atgatcctcc	agcgcgggga	tctcatgctg	gagttctctg	cccacccggg	atcgatocaa	12300
cacttacgtt	tgcaacgtcc	aagagcaaat	agaccacgaa	cgccggaagg	ttgccgcagc	12360
gtgtggattg	cgtctcaatt	ctctcttgca	ggaatgcaat	gatgaatatg	atactgacta	12420
tgaaactttg	aggaataact	gcctagcacc	gtcacctcat	aacgtgcate	atgcatgccc	12480
tgacaacatg	gaacatcgct	atTTTTtga	agaattatgc	tcgttgaggg	atgtcgcggc	12540
aattgcagct	attgccaaca	tcgaaactacc	cctcacgcat	gcattcatca	atattattca	12600
tgccgggaaa	ggcaagatta	atccaaactgg	caaatcatcc	agcgtgattg	gtaacttcag	12660
ttccagcgac	ttgattcgtt	ttggtgctac	ccacgttttc	aataaggacg	agatgggtgga	12720
gtaaagaagg	agtgcgctga	agcagatcgt	tcaaacattt	ggcaataaag	tttcttaaga	12780
ttgaatcctg	ttgcccgtct	tcgcatgatt	atcatataat	ttctggtgaa	ttacgttaag	12840
catgtaataa	ttaacatgta	atgcatgacg	ttatttatga	gatgggtttt	tatgattaga	12900
gtcccgcaat	tatacattta	atcgcgata	gaaaacaaaa	tatagcgcgc	aaactaggat	12960
aaattatcgc	gcgcgggtgc	atctatgtta	ctagatcgat	caaacttcgg	tactgtgtaa	13020
tgacgatgag	caatcgagag	gctgactaac	aaaaggtaca	tcgcgatgga	tcgatccatt	13080
cgccattcag	gctgcgcaac	tgttgggaag	ggcgatcggg	gcgggcctct	tcgctattac	13140
gccagctggc	gaaaggggga	tgtgctgcaa	ggcgattaag	ttgggtaacg	ccagggtttt	13200
cccagtcacg	acgttgtaaa	acgacggcca	gtgaattcct	gcagcccggg	ggatccgccc	13260
actcgaggcg	cgccgtcgac	ggatataatg	agccgtaaac	aaagatgatt	aagtagtaat	13320
taatacgtac	tagtaaaagt	ggcaaaagat	aacgagaaaag	aaccaatttc	tttgcattcg	13380
gccttagcgg	aaggcatata	taagcttga	ttattttatt	tagtgtaatg	atttcgtaca	13440

-continued

---

accaaagcat	ttatntagta	ctctcacact	tgtgtcgcgg	cggccgcta	caggaacag	13500
tggtggcggc	cctcggcgcg	ctcgtactgc	tccacgatgg	tgtagtcttc	gttgtgggag	13560
gtgatgtcca	gcttgaggtc	cacgtagtag	tagccgggca	gctgcacggg	cttcttgccc	13620
atgtagatgg	acttgaactc	caccaggtag	tgcccgccgt	ccttcagctt	cagggccttg	13680
tggatctcgc	ccttcagcac	gccgtcgcgg	gggtacaggc	gctcggtgga	ggcctcccag	13740
cccatagtct	tctctcgcac	tacggggccg	tcggagggga	agttcacgcc	gatgaacttc	13800
accttgtaga	tgaaggagcc	gtcctgcagg	gaggagtctc	gggtcacggg	caccacgccg	13860
ccgtcctcga	agttcatcac	gcgctccac	ttgaagccct	cggggaagga	cagcttcttg	13920
tagtcgggga	tgctggcggg	gtgcttcacg	tacaccttgg	agccgtactg	gaactggggg	13980
gacaggatgt	cccaggcga	gggcaggggg	ccgccttgg	tcacctcag	cttggcggtc	14040
tgggtgcctc	cgtagggcgc	gccctcgcgc	tcgccctcga	tctcgaactc	gtggccgctc	14100
acggagccct	ccatgcgcac	cttgaagcgc	atgaactcct	tgatgacgtc	ctcggaggag	14160
gccatgggccc	gcttgggggg	ctatggaaga	ctttcttagt	tagttgtgtg	aataagcaat	14220
gttgggagaa	tcgggactac	ttataggata	ggaataaaac	agaaaagtat	taagtgctaa	14280
tgaatatatt	agactgataa	ttaaaatctt	cacgtatgtc	cacttgatata	aaaaacgtca	14340
ggaataaagg	aagtacagta	gaatttaaag	gtactctttt	tatatatacc	cgtgtctctc	14400
ttttggctag	ctagttgcat	aaaaataat	ctatatTTTT	atcattattt	taaatactct	14460
atgagatggt	aaatatttat	cataattttt	tttactatta	tttattattt	gtgtgtgtaa	14520
tacatataga	agttaattac	aaattttatt	tactttttca	ttattttgat	atgattcacc	14580
attaatttag	tgttattatt	tataatagtt	cattttaatc	ttttgtata	tattatgcgt	14640
gcagtacttt	tttctacat	ataactacta	ttacatttta	tttatataat	atttttatta	14700
atgaattttc	gtgataatat	gtaatattgt	tcattattat	ttcagatttt	ttaaaaatat	14760
ttgtgttatt	atztatgaaa	tatgtaattt	tttttagtatt	tgattttatg	atgataaagt	14820
gttctaaatt	caaaaagagg	gggaaagcgt	aaacattaaa	aaacgctatc	aaacaaaaac	14880
aaaatcttgt	taataaagat	aaaactggtt	gttttgatca	ctgttatttc	gtaataataa	14940
aacattattt	atatttatat	tgttgacaac	caaatttgcc	tatcaaatct	aaccaatata	15000
atgcatgcgt	ggcaggtaat	gtactaccat	gaacttaagt	catgacataa	taaacctgta	15060
atctgaccaa	tgcatgtacc	tanctaaatt	gtattttgta	cacgaagcaa	atgattcaat	15120
tcacaatgga	gatgggaaac	aaataatgaa	gaaccagaaa	ctaagaaagc	tttctgaaa	15180
aataaaataa	aggcaatgtc	aaaagtatac	tgcatcatca	gtccagaaag	cacatgatata	15240
ttttttatca	gtatcaatgc	agctagtttt	attttacaat	atcgatatag	ctagtttaaa	15300
tatattgcag	ctagatttat	aaatatttgt	gttattattt	atcatttggtg	taatcctggt	15360
tttagtattt	tagtttatata	atgatgataa	tgtattccaa	atttaaaaga	agggaaataa	15420
atttaaacaa	gaaaaaaagt	catcaaacaa	aaaacaaatg	aaaggtgga	aagatgttac	15480
catgtaatgt	gaatgttaca	gtatttcttt	tattatagag	ttaacaaatt	aactaatatg	15540
attttgtaa	taatgataaa	atattttttt	tattattatt	tcataatata	aaaaatagttt	15600
acttaatata	aaaaaaatc	tatcgttcac	aacaaagttg	gccacctaat	ttaacctgctc	15660
atgtacccat	ggaccatatt	aggtaacct	caaacctgat	gaagagataa	agagatgaag	15720
acttaagtca	taacacaaaa	ccataaaaaa	caaaaataca	atcaaccgctc	aatctgacca	15780

-continued

---

```

atgcatgaaa aagctgcaat agtgagtggc gacacaaagc acatgatttt cttacaacgg 15840
agataaaaacc aaaaaaatat ttcatgaaca acctagaaca aataaagctt ttatataata 15900
aatatataaa taaataaagg ctatggaata atataacttca atatatttgg attaaataaa 15960
ttgttgccgg ggttgatata tttatacaca cctaaagtca cttcaatctc attttcactt 16020
aacttttatt ttttttttct ttttatttat cataaagaga atattgataa tataacttttt 16080
aacatatttt tatgacattt tttattggtg aaaacttatt aaaaatcata aattttgtaa 16140
gttagattta tttaaagagt tcctcttctt attttaaatt ttttaataaa tttttaaata 16200
actaaaattt gtgttaaaaa tgttaaaaaa gtgtgttatt aaccttctc ttcgaggatc 16260
cgtaccgagc tcggatctc tagaaatccg tcaacatggt ggagcacgac actctcgtct 16320
actccaagaa tatcaaagat acagtctcag aagaccaaag ggctattgag acttttcaac 16380
aaagggtaat atcgggaaac ctctcggat tccattgccc agctatctgt cacttcatca 16440
aaaggacagt agaaaaggaa ggtggcacct acaaatgcca tcattgcatg aaaggaaagg 16500
ctatcgttca agatgcctct gccgacagt gtcccaaga tggaccccca cccacgagga 16560
gcatcgtgga aaaagaagac gttccaacca cgtcttcaaa gcaagtggat tgatgtgatg 16620
atcctatgcg tatggtatga cgtgtgttca agatgatgac ttcaaaccta cctatgacgt 16680
atggtatgaa cgtgtgtoga ctgatgactt agatcactc gagcggctat aaatacgtac 16740
ctacgcaccc tgcgctacca tcctagagc tgcagcttat tttacaaca attaccaaca 16800
acaacaaaca acaacaaaca ttacaattac tatttacaat tacagtcgac ccgggatcgt 16860
acctctaggg tggcgccgc atggagagat ctcaacggca gtctcctccg ccaccgtcgc 16920
cgtcctcctc ctgcctcctc gtcctccgcg acaccgtcct cgctcctccc ggaagaggc 16980
ggagggcgcc gacggccaag gccggcgccg agcctaataa gaggatccgc aaggaccccg 17040
ccgccgccgc cgcggggaag aggagctccg tctacagggg agtcaccagg cacaggtgga 17100
cgggcaggtt cgaggcgcat ctctgggaca agcaactgct cgccgcgctc cacacaaga 17160
agaaaggcag gcaagtctac ctgggggct atgacagcga ggaggcagct gctcgtgcct 17220
atgacctcgc agctctcaag tactggggtc ctgagactct gctcaacttc cctgtggagg 17280
attactccag cgagatgccg gagatggagg ccgtgtcccg ggaggagtac ctggcctccc 17340
tccgccgcag gagcagcggc ttctccaggg gcgtctccaa gtacagagge gtcgccagge 17400
atcaccacaa cgggaggtgg gagcacgga ttggcgagt ctttgggaac aagtacctct 17460
acttgggaac atttgacact caagaagagg cagccaagge ctatgacctt gcggccattg 17520
aataaccgtg cgtcaatgct gtaaccaact tcgacatcag ctgctacctg gaccaccgce 17580
tgttcctggc acagctccaa caggagccac aggtggtgcc ggcaactaac caagaacctc 17640
aacctgatca gagcgaaacc ggaactacag agcaagagcc ggagtcaagc gaagccaaga 17700
caccggatgg cagtgcagaa cccgatgaga acgcggtgcc tgacgacacc gcggagcccc 17760
tcaccacagt cgacgacagc atcgaagagg gcttgtggag cccttgcag gattacgagc 17820
tagacaccat gtcgagacca aactttggca gctcaatcaa tctgagcgag tggttcgctg 17880
acgcagactt cgactgcaac atcggatgcc tgttcgatgg gtgttctgcg gctgacgaag 17940
gaagcaagga tgggtgaggt ctggcagatt tcagtctgtt tgaggcaggt gatgtccagc 18000
tgaaggatgt tctttcggat atggaagagg ggatacaacc tccagcgatg atcagtgtgt 18060
gcaacgcggc cgcaagtatg aactaaaatg catgtagggt taagagctca tggagagcat 18120
ggaatattgt atccgacct gtaacagat aataactgag ctccatctca cttctctat 18180

```

-continued

---

```

gaataaacaa aggatgttat gatataataa cactctatct atgcacctta ttgttctatg 18240
ataaatttcc tcttattatt ataatcctc tgaatcgtga cggcttatgg aatgcttcaa 18300
atagtacaaa aacaaatgtg tactataaga ctttctaacc aattctaacc ttagcattgt 18360
gaacgagaca taagtgttaa gaagacataa caattataat ggaagaagtt tgtctccatt 18420
tatatattat atattaccca cttatgtatt atattaggat gttaaggaga cataacaatt 18480
ataaagagag aagtttgtat ccatttatat attatatact acccatttat atattatact 18540
tatccactta tttaatgtct ttataagggt tgatccatga tatttctaatt attttagttg 18600
atatgtatat gaaagggtag tatttgaact ctcttactct gtataaagggt tggatcatcc 18660
ttaaagtggt tctatttaatt tttattgctt cttacagata aaaaaaaaaat tatgagttgg 18720
tttgataaaa tattgaagga tttaaaataa taataaataa catataatat atgtatataa 18780
atattattata atataacatt tatctataaa aaagtaaata ttgtcataaa tctatacaat 18840
cgtttagcct tgctggacga atctcaatta tttaaacgag agtaaacata tttgactttt 18900
tggttattta acaaatattt atttaacact atatgaaatt ttttttttta tcagcaaaga 18960
ataaaattaa attaagaagg acaatgggtg cccaatcctt atacaaccaa cttccacaag 19020
aaagtcaagt cagagacaac aaaaaaacaa gcaaaggaaa ttttttaatt tgagttgtct 19080
tgtttctgctc ataatttatg cagtaaaaca ctacacataa cccttttagc agtagagcaa 19140
tggttgaccg tgtgcttagc ttcttttatt ttattttttt atcagcaaag aataaataaa 19200
ataaaatgag acacttcagg gatgtttcaa caagctctag agggcccaat tcgccctata 19260
gtgagtcgta ttacaattca ctggccgtcg ttttacaacg tcgtgactgg gaaaacctg 19320
gcgttaccca acttaatcgc cttgcagcac atcccccttt cgcagctgg cgtaatagcg 19380
aagaggcccg caccgatcgc ccttcccaac agttgcgcag cctatacgtg cgagatcccg 19440
cgggccagat cctgcaggag atccaagctt gg 19472

```

&lt;210&gt; SEQ ID NO 36

&lt;211&gt; LENGTH: 1188

&lt;212&gt; TYPE: DNA

&lt;213&gt; ORGANISM: Zea mays

&lt;400&gt; SEQUENCE: 36

```

atggagagat ctcaacggca gtctcctcgg ccaccgtcgc cgtcctctc ctcgtctccc 60
gtctccgcgg acaccgtcct cgtccctccc ggaaagaggg ggagggcggc gacggccaag 120
gccggcgcgg agcctaataa gaggatccgc aaggaccccg ccgcccgcgc cgcggggaag 180
aggagctccg tctacagggg agtcaccagg cacaggtgga cgggcagggt cgaggcgcat 240
ctctgggaca agcactgcct cgcgcgcgtc cacaacaaga agaaaggcag gcaagtctac 300
ctggggcgt atgacagcga ggaggcagct gctcgtgcct atgacctcgc agctctcaag 360
tactggggtc ctgagactct gctcaacttc cctgtggagg attactccag cgagatgccg 420
gagatggagg ccgtgtcccg ggaggagtac ctggcctccc tccgcccag gagcagcggc 480
ttctccaggg gcgtctccaa gtacagaggc gtcgccaggg atcaccacaa cgggaggtgg 540
gaggcacgga ttgggcgagt ctttgggaac aagtacctct acttgggaac atttgacact 600
caagaagagg cagccaaggg ctatgacctt gcggccattg aataccgtgg cgtaaatgct 660
gtaaccaact tcgacatcag ctgctacctg gaccaccgcg tgttctctggc acagctccaa 720
caggagccac aggtggtgdc ggcactcaac caagaacctc aacctgatca gagcgaacc 780

```



-continued

---

Leu Trp Ser Pro Cys Met Asp Tyr Glu Leu Asp Thr Met Ser Arg Pro  
 305 310 315 320

Asn Phe Gly Ser Ser Ile Asn Leu Ser Glu Trp Phe Ala Asp Ala Asp  
 325 330 335

Phe Asp Cys Asn Ile Gly Cys Leu Phe Asp Gly Cys Ser Ala Ala Asp  
 340 345 350

Glu Gly Ser Lys Asp Gly Val Gly Leu Ala Asp Phe Ser Leu Phe Glu  
 355 360 365

Ala Gly Asp Val Gln Leu Lys Asp Val Leu Ser Asp Met Glu Glu Gly  
 370 375 380

Ile Gln Pro Pro Ala Met Ile Ser Val Cys Asn  
 385 390 395

<210> SEQ ID NO 38  
 <211> LENGTH: 1239  
 <212> TYPE: DNA  
 <213> ORGANISM: Glycine max

<400> SEQUENCE: 38

atgaagaggt ctccagcatc ttcttgttca tcatctactt cctctgttgg gtttgaagct 60  
 cccattgaaa aaagaaggcc taagcatcca aggaggaata atttgaagtc acaaaaatgc 120  
 aagcagaacc aaaccaccac tgggtggcaga agaagctcta tctatagagg agttacaagg 180  
 cataggtgga cagggaggtt tgaagctcac ctatgggata agagctcttg gaacaacatt 240  
 cagagcaaga agggctgaca agtttatttg ggggcatatg atactgaaga atctgcagcc 300  
 cgtacctatg accttgcagc ccttaaacat tggggaaaag atgcaacctt gaatttcccg 360  
 atagaaaact ataccaagga gctcagaggaa atggacaagg tttcaagaga agaataattg 420  
 gcttctttgc ggcgccaaag cagtggcttt tctagaggcc tgtctaagta ccgtgggggt 480  
 gctaggcatc atcataatgg tcgctgggaa gcacgaattg gaagagtatg cggaacaag 540  
 tacctctact tggggacata taaaactcaa gaggaggcag cagtggcata tgacatggca 600  
 gcaatagagt accgtggagt caatgcagtg accaattttg acataagcaa ctacatggac 660  
 aaaataaaga agaaaaatga ccaaacccaa caacaacaaa cagaagcaca aacggaaca 720  
 gttcctaact cctctgactc tgaagaagta gaagtagaac aacagacaac aacaataacc 780  
 acaccacccc catctgaaaa tctgcacatg ccaccacagc agcaccaagt tcaatacacc 840  
 ccccatgtct ctccaagga agaagaatca tcatcactga tcacaattat ggaccatgtg 900  
 cttgagcagg atctgccatg gagcttcatg tacactggct tgtctcagtt tcaagatcca 960  
 aacttggett tctgcaaagg tgatgatgac ttggtgggca tgtttgatag tgcaggggtt 1020  
 gaggaagaca ttgattttct gttcagcact caacctgggtg atgagactga gactgatgtc 1080  
 aacaatatga ggcagtttt ggatagtgtt gagtgtggag acacaaatgg ggctggtgga 1140  
 agcatgatgc atgtggataa caagcagaag atagtatcat ttgcttcttc accatcatct 1200  
 acaactacag tttcttgtga ctatgctcta gatctatga 1239

<210> SEQ ID NO 39  
 <211> LENGTH: 412  
 <212> TYPE: PRT  
 <213> ORGANISM: Glycine max

<400> SEQUENCE: 39

Met Lys Arg Ser Pro Ala Ser Ser Cys Ser Ser Ser Thr Ser Ser Val  
 1 5 10 15



-continued

&lt;213&gt; ORGANISM: Momordica charantia

&lt;400&gt; SEQUENCE: 40

```

atgagaaggt ctcctctgt ttctacttcc tctctctcct cctcctcctg cgtcggcggc   60
ggcggcttcg acagcaataa tctcaatctc gccgcccctc cgcgcgcggc gcaatcggag   120
aagaccggag cgaaaccccg gaagcgggaat caggacgacg ccaaatgcga gattgagaat   180
cgtaacggta ataacaacaa cagcagcaac aacaatgcct cttccggccg cgggagctcc   240
atttacagag gagtacttag gcaccgatgg accggccggt tcgaagcgcg tctctgggac   300
aagagtctcg ggaatagcat tcagaacaaa aaaggaaggc aagtttattt gggagcatac   360
gataacgagg aagctgcccg ccgaacttat gacctcctg ccctcaagta ctgggggtccc   420
ggaaccaccc tcaatttccc ggtagagctg tacaggaatg aaatagaaga aatgcggaaa   480
gttacgaagg aggagtattt ggcgctgcta cggcggcggg gcagcggatt ttcgagaggc   540
gtatcgaagt accgcccgtg ggcgccccac caccacaacg gccgggtggg ggcgcgggat   600
ggccgtgttt tcggaagcaa atatctttac ctgggaactt acaacacaca agaggaagca   660
gcagcagcat atgacatggc tgcaattgag tacagagggg tcaatgcagt gaccaatttc   720
gacatcagca attacattgg gcggtgggag aataaatcat cagtttttcc agcagcagag   780
cagcccctac agcccactg ctcccctget tcctcttctg aggaaggcga agtagtacag   840
cagcaacagc aacagacgac gatggcgttc tcaggctcgc ccctccagtt cccgtcgatg   900
gagaacagcc cgacgacaat ggaggaggat catgatctgc attggtcatt cctagacacg   960
gggttcctgc aggtcccoga cctcccctc gagaaagtct gcgaattgcc tgacctgttc  1020
tttgatgaga tcgggttcga ggacgacatc gggttgatat tcgagggcag cttggaagac  1080
gagaggtgcg gggagggggg tgagaagtta gaagatgtgg ggaaaatgga gatgatgaag  1140
agtgatcatg aggagagggg gttgttctcg actacttcgc catcttcgct gtcgataacc  1200
acctcggttt cgtgtgaatt tagggtttga  1230

```

&lt;210&gt; SEQ ID NO 41

&lt;211&gt; LENGTH: 409

&lt;212&gt; TYPE: PRT

&lt;213&gt; ORGANISM: Momordica charantia

&lt;400&gt; SEQUENCE: 41

```

Met Arg Arg Ser Pro Ser Val Ser Thr Ser Ser Ser Ser Ser Ser Ser
 1           5              10              15

Cys Val Gly Gly Gly Gly Phe Asp Ser Asn Asn Leu Asn Leu Ala Ala
 20              25              30

Pro Pro Arg Arg Pro Gln Ser Glu Lys Thr Gly Ala Lys Arg Arg Lys
 35              40              45

Arg Asn Gln Asp Asp Ala Lys Cys Glu Ile Glu Asn Arg Asn Gly Asn
 50              55              60

Asn Asn Asn Ser Ser Asn Asn Asn Ala Ser Ser Gly Arg Arg Ser Ser
 65              70              75              80

Ile Tyr Arg Gly Val Thr Arg His Arg Trp Thr Gly Arg Phe Glu Ala
 85              90              95

His Leu Trp Asp Lys Ser Ser Trp Asn Ser Ile Gln Asn Lys Lys Gly
 100             105             110

Arg Gln Val Tyr Leu Gly Ala Tyr Asp Asn Glu Glu Ala Ala Ala Arg
 115             120             125

Thr Tyr Asp Leu Ala Ala Leu Lys Tyr Trp Gly Pro Gly Thr Thr Leu

```

-continued

130	135	140																		
Asn	Phe	Pro	Val	Glu	Ser	Tyr	Arg	Asn	Glu	Ile	Glu	Glu	Met	Arg	Lys					
145					150					155					160					
Val	Thr	Lys	Glu	Glu	Tyr	Leu	Ala	Ser	Leu	Arg	Arg	Arg	Ser	Ser	Gly					
				165						170					175					
Phe	Ser	Arg	Gly	Val	Ser	Lys	Tyr	Arg	Gly	Val	Ala	Arg	His	His	His					
			180						185					190						
Asn	Gly	Arg	Trp	Glu	Ala	Arg	Ile	Gly	Arg	Val	Phe	Gly	Ser	Lys	Tyr					
		195						200					205							
Leu	Tyr	Leu	Gly	Thr	Tyr	Asn	Thr	Gln	Glu	Glu	Ala	Ala	Ala	Ala	Tyr					
	210					215					220									
Asp	Met	Ala	Ala	Ile	Glu	Tyr	Arg	Gly	Val	Asn	Ala	Val	Thr	Asn	Phe					
225					230						235				240					
Asp	Ile	Ser	Asn	Tyr	Ile	Gly	Arg	Leu	Glu	Asn	Lys	Ser	Ser	Val	Phe					
				245						250					255					
Pro	Ala	Ala	Glu	Gln	Pro	Leu	Gln	Pro	Asn	Cys	Ser	Pro	Ala	Ser	Ser					
			260						265						270					
Ser	Glu	Glu	Gly	Glu	Val	Val	Gln	Gln	Gln	Gln	Gln	Gln	Thr	Thr	Met					
		275						280						285						
Ala	Phe	Ser	Gly	Ser	Pro	Leu	Gln	Phe	Pro	Ser	Met	Glu	Asn	Ser	Pro					
290						295					300									
Thr	Thr	Met	Glu	Glu	Asp	His	Asp	Leu	His	Trp	Ser	Phe	Leu	Asp	Thr					
305					310						315				320					
Gly	Phe	Val	Gln	Val	Pro	Asp	Leu	Pro	Leu	Glu	Lys	Ser	Gly	Glu	Leu					
				325						330					335					
Pro	Asp	Leu	Phe	Phe	Asp	Glu	Ile	Gly	Phe	Glu	Asp	Asp	Ile	Gly	Leu					
			340						345					350						
Ile	Phe	Glu	Ala	Ser	Leu	Glu	Asp	Glu	Arg	Cys	Gly	Glu	Gly	Gly	Glu					
		355					360						365							
Lys	Leu	Glu	Asp	Val	Gly	Lys	Met	Glu	Met	Met	Lys	Ser	Asp	His	Glu					
370						375						380								
Glu	Arg	Gly	Leu	Phe	Ser	Thr	Thr	Ser	Pro	Ser	Ser	Ser	Ser	Ile	Thr					
385					390					395					400					
Thr	Ser	Val	Ser	Cys	Glu	Phe	Arg	Val												
				405																

&lt;210&gt; SEQ ID NO 42

&lt;211&gt; LENGTH: 430

&lt;212&gt; TYPE: PRT

&lt;213&gt; ORGANISM: Arabidopsis thaliana

&lt;400&gt; SEQUENCE: 42

Met	Lys	Lys	Arg	Leu	Thr	Thr	Ser	Thr	Cys	Ser	Ser	Ser	Pro	Ser	Ser					
1				5					10					15						
Ser	Val	Ser	Ser	Ser	Thr	Thr	Thr	Ser	Ser	Pro	Ile	Gln	Ser	Glu	Ala					
				20					25					30						
Pro	Arg	Pro	Lys	Arg	Ala	Lys	Arg	Ala	Lys	Lys	Ser	Ser	Pro	Ser	Gly					
			35					40						45						
Asp	Lys	Ser	His	Asn	Pro	Thr	Ser	Pro	Ala	Ser	Thr	Arg	Arg	Ser	Ser					
		50				55							60							
Ile	Tyr	Arg	Gly	Val	Thr	Arg	His	Arg	Trp	Thr	Gly	Arg	Phe	Glu	Ala					
65					70					75					80					
His	Leu	Trp	Asp	Lys	Ser	Ser	Trp	Asn	Ser	Ile	Gln	Asn	Lys	Lys	Gly					
				85					90						95					

-continued

---

Lys Gln Val Tyr Leu Gly Ala Tyr Asp Ser Glu Glu Ala Ala Ala His  
 100 105 110

Thr Tyr Asp Leu Ala Ala Leu Lys Tyr Trp Gly Pro Asp Thr Ile Leu  
 115 120 125

Asn Phe Pro Ala Glu Thr Tyr Thr Lys Glu Leu Glu Glu Met Gln Arg  
 130 135 140

Val Thr Lys Glu Glu Tyr Leu Ala Ser Leu Arg Arg Gln Ser Ser Gly  
 145 150 155 160

Phe Ser Arg Gly Val Ser Lys Tyr Arg Gly Val Ala Arg His His His  
 165 170 175

Asn Gly Arg Trp Glu Ala Arg Ile Gly Arg Val Phe Gly Asn Lys Tyr  
 180 185 190

Leu Tyr Leu Gly Thr Tyr Asn Thr Gln Glu Glu Ala Ala Ala Tyr  
 195 200 205

Asp Met Ala Ala Ile Glu Tyr Arg Gly Ala Asn Ala Val Thr Asn Phe  
 210 215 220

Asp Ile Ser Asn Tyr Ile Asp Arg Leu Lys Lys Lys Gly Val Phe Pro  
 225 230 235 240

Phe Pro Val Asn Gln Ala Asn His Gln Glu Gly Ile Leu Val Glu Ala  
 245 250 255

Lys Gln Glu Val Glu Thr Arg Glu Ala Lys Glu Glu Pro Arg Glu Glu  
 260 265 270

Val Lys Gln Gln Tyr Val Glu Glu Pro Pro Gln Glu Glu Glu Lys  
 275 280 285

Glu Glu Glu Lys Ala Glu Gln Gln Glu Ala Glu Ile Val Gly Tyr Ser  
 290 295 300

Glu Glu Ala Ala Val Val Asn Cys Cys Ile Asp Ser Ser Thr Ile Met  
 305 310 315 320

Glu Met Asp Arg Cys Gly Asp Asn Asn Glu Leu Ala Trp Asn Phe Cys  
 325 330 335

Met Met Asp Thr Gly Phe Ser Pro Phe Leu Thr Asp Gln Asn Leu Ala  
 340 345 350

Asn Glu Asn Pro Ile Glu Tyr Pro Glu Leu Phe Asn Glu Leu Ala Phe  
 355 360 365

Glu Asp Asn Ile Asp Phe Met Phe Asp Asp Gly Lys His Glu Cys Leu  
 370 375 380

Asn Leu Glu Asn Leu Asp Cys Cys Val Val Gly Arg Glu Ser Pro Pro  
 385 390 395 400

Ser Ser Ser Ser Pro Leu Ser Cys Leu Ser Thr Asp Ser Ala Ser Ser  
 405 410 415

Thr Thr Thr Thr Thr Thr Ser Val Ser Cys Asn Tyr Leu Val  
 420 425 430

<210> SEQ ID NO 43  
 <211> LENGTH: 2004  
 <212> TYPE: DNA  
 <213> ORGANISM: Arabidopsis thaliana

<400> SEQUENCE: 43

```

gggtactctt tacatggtc tttactccgt ctcaaaattt cctttttttg ttggctctct    60
ccgaacgagt tggagaaatc gttaacccta atcgaagatc tagattcctc tacatacgtt    120
tgatctctct ctcagtatgg attacaaagc gccaaaggaga tactactcac acggagtgtg    180
tgcgagacag caagatttcg caacagatat agttacgaga agaagacctt atgtccctta    240
    
```

-continued

---

```

cgaccgtcca aataagtttt caaggagtct ggtttggacg tcaaaagagt acaaatcacc 300
cgagggcaat aatatgccaa ggaccaatga tgtgtcaccg aaaccaccag ttttaggttt 360
ggcgaggaag aatgctgctt gtgggccaat gagatcttct agtctcagaa aatgggtatg 420
taagtattgg aaagatggaa agtgcaagag gggtgagcag tgccagtctt tacactcttg 480
gtcttgtttc cctggattgg ccatggtagc ttctctttaa gggcacaata aggaactaaa 540
ggggatcgct ctccctgagg gttcagataa actcttttca gtcagtattg atggtagatt 600
gcgagtttgg gactgcaatt ctggtcagtg tgtacattcc atcaaccttg acgcagaagc 660
agggctctcta atcagtgaag gcccttgggt tttccttggc ttgccaaacg ctataaaggc 720
ttttaacggt caaaccagtc aagatttga tcttcaagca gcaggggtgg ttggtcaggt 780
gaatgcaatg actattgcaa acggaatgct ttttgcctga acaagttctg gtagtatctt 840
agtctgaaa gctactacag actctgagtc tgatccattc aaactctga catctcttga 900
gggacatagt ggtgaagtca cttgttttgc tgttggaggt caaatgctat actctggttc 960
tgtcgataaa acaatcaaga tgtgggatct caacacctg caatgataa tgacctgaa 1020
gcaacatacc ggcactgtca cttcactctt atgttgggat aaatgtttga tatcgtcttc 1080
cttgatggg accataaaag tttgggctta ttctgaaaac ggaatcttga aagttgttca 1140
aactcgcaga caagaacaga gtagtgttca tgctcttctt ggtatgcatg atgcagaagc 1200
caaaccgata atattctgct cttacaaaaa cggaaccggt ggcattttcg acctaccatc 1260
ttttcaagaa agaggaagga tgttctctac gcacacgac gccacactca caattggctc 1320
tcaaggattg ttattcagtg gagacgagag tggtaacttg cgtgtatgga ccttagctgc 1380
tggcaacaaa gtttagtctt ttgcactaaa gaattctgat ttaattttgt ggtttatatg 1440
ttgagttaac tgtaagaga gttttatttt gtaatagggt tatcagtc aaacaatct 1500
ttgtatcaac caaatgtaat ttttctggtt aattcgattt cagagttttt actttaagat 1560
aaacaaactc tttcacacat catttaatga aagtgagaaa gcttaaaaaa caaacaaaga 1620
aactgatcca tttttggcgg gtctctctct actcttattc atatgtgtta acgaactata 1680
gcgtaaaatt cagagcaagc gatctccgat ttgaacgtgg ctatcaccgg aggcccacca 1740
ctacggcgca tacgctctaa gtgaggatta aagtgcctcg gtggtgacgt tgaagaaact 1800
cgcccatggt ttttgttate tctgcagcca agtgtcgttc tttcttcgcc acttctcttc 1860
aagctacagt gaatttaaaa atggcgtctt tctttgatct cgtatacata agctggattg 1920
gtttcttaaa caaattctct tcttttggg tcttctgggt ttgccttgta agtgtttgtg 1980
tttttgcctc tgagaaaaaa tcgc 2004

```

```

<210> SEQ ID NO 44
<211> LENGTH: 2790
<212> TYPE: DNA
<213> ORGANISM: Brassica napus
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: (2776)..(2776)
<223> OTHER INFORMATION: n is a, c, g, or t

```

```

<400> SEQUENCE: 44

```

```

ggccacttct catcatgtta cagggaccat aaaaatggcg tatttcttca gccccgggta 60
taaatacaca catgatctg tgggtgttts ttccacaagt tacatctctt tctggttttt 120
gtattgcaag tgtttgtgtt ttttgcctcc gagagaaaaa catgccgacc ggtaggttcg 180
agacgatgog tgaatgggtc cagcagcga tctctgctca acgcaatgag ctctctctc 240

```

-continued

---

ttttttccag atacgtagct caggggaaaag ggatactgya gteccaccag ctgattgacg	300
agtccctcaa gactgtgaaa gtggatggaa ctacagaaga tcttaagaat cgtcccttca	360
tgaaagttct gcagctctgca gaggaagcca tagttttgcc tccctttggt gcscctggcga	420
ttcgtcccag acctggtggt agagaatatg tccgtgtgaa tgtctacgag ctgagcgtag	480
accatttaac tgtttctgag tatcttcggt tcaaggaaga gctcgttaat ggccatgcca	540
atgggaatta tctctcagag cttgattttg aaccgttcaa cgcaacgttt cctcgtccaa	600
ctcggtcac c atctattggg aatgggggttc agttcctcaa ccgtcacctc tcgtcaatca	660
tgttccgtaa caaagacagc ttggagcctt tgcttgagtt tctccgact cacaacatg	720
acggccgtgc catgatgctg aatgatcgaa tacagaacat ccgcacactt caggaagctt	780
tggcgagggc agaggagttc ctctctaaac ttcctttggc tacaccatac tctgaattcg	840
aatttgract acaagggatg ggatttgaga ggggatgggg tgacacgkca cagaaggttt	900
cagaaatggt gcactctmctt ctggacatac tccaggcacc tgatccttct gtcttgagga	960
cgtttcttgg aaggattcct atgggtttca atgtygkat tttgtctccg catggctact	1020
ttggccaagc caatgtcttg ggtcttctg atactggtgg acaggttgc tacattcttg	1080
atcaagtacg tgctttggaa agcgagatgc tctyaggat acagaagcaa ggactggatg	1140
ttactccaaa gattctcatt gtaacaaggt tgataccaga agcagaagga acaacatgca	1200
accagaggtt agaaaargtw agcggtagc aacacrcaca tattctrcga atacctttm	1260
ggactgaaaa gggcattcct cgcaagtgga tctcgaggtt tgatgtctgg ccatacctgg	1320
agactttcgc agaggatgca tcaaatgaaa ttgctgcgga gttgcaaggt gtgccaatc	1380
tcatcattgg caactacagt gatgggaatc tcgtggcttc tttgttagct tgtaagctag	1440
gcgtgataca gtgcaatatt gctcatgctt tggagaaaac caagtatcca gagtctgaca	1500
tttactggag aaaccatgaa gataagtatc attttgcaag tcagttcact gcggaactaa	1560
ttgccatgaa taatgctgat ttcacatca ccagcacata ccaagagatc gctggaagca	1620
aaaacaaagt tgggcaatac garagccaca cagctttcac ccttctgggt ctttacagag	1680
ttgtkcatgg aatcaatgct tttgatocca agtttaatat agtctctcca ggagctgata	1740
tgaccatata ctyccwat tctgacaagg aaagaagact aactgccctt catgagtcwa	1800
ttgaagaact yctgtttagc agygaacaga atgttgagca tggttggttt ctkagegacc	1860
agwyaagcc aatcatttcc tccatggcca gacttgacag agtgaaaaac ttgactgggc	1920
tagttgagtg ctatgccaa aacrgcaasc tgagagaggt tgccaaccty sttgastwg	1980
gtggctacgt ggacgtgaat cagtccaggg acagagagga aatggctgag atacaaaaga	2040
tgcacagcct ratcaagcag tatggtttac acgggtgagtt caggtggata gctgctcaa	2100
tgaaccgtgc tmggaacgggt gagctttacc gttatatcgc agacacwaaa ggtgttttg	2160
ttcagcctgc tttctatgaa gcktttgggc tcacagttgt ggaatcaatg acttggtggc	2220
tcccaacggt tgctacatgt catggtggac ctgaggagat catcgagaat ggagtttctg	2280
gcttccacat cgaccctwat catccagaac agsttgcaac tactttggtc agcttcttyg	2340
agacctgcaa cgctgatcca agtcaactgg agaaaaatctc tgatggaggg ctttaagcgaa	2400
tctatgaaag gtacacatgg aagaagtact cagagaggtc gcttacgctg gctggtgtct	2460
attcattctg gaaacatggt tctaagcttg aaaggagaga aacacgacgt tacctagaga	2520
tgtttactc tctcaagtat cgtgatctgg ccaattcaat cccactggca actgatgagc	2580

-continued

---

```

attgagcaag ctatggttgg attctaatac ttgctgcact ccctggttg tgtttctgtt 2640
atctttgaat aaataagcta ttgctggctt ttgtttccat gactagtttg gttttcagac 2700
ttttctgtt gttttcttga tatgaataac aagtatcggtt gagttctaag ctgggcatta 2760
aataacttgt cgtgtnggaa agcttactga 2790

```

```

<210> SEQ ID NO 45
<211> LENGTH: 807
<212> TYPE: PRT
<213> ORGANISM: Brassica napus
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: (40)..(40)
<223> OTHER INFORMATION: Xaa can be any naturally occurring amino acid
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: (229)..(229)
<223> OTHER INFORMATION: Xaa can be any naturally occurring amino acid
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: (243)..(243)
<223> OTHER INFORMATION: Xaa can be any naturally occurring amino acid
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: (359)..(359)
<223> OTHER INFORMATION: Xaa can be any naturally occurring amino acid
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: (568)..(568)
<223> OTHER INFORMATION: Xaa can be any naturally occurring amino acid
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: (593)..(593)
<223> OTHER INFORMATION: Xaa can be any naturally occurring amino acid
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: (595)..(596)
<223> OTHER INFORMATION: Xaa can be any naturally occurring amino acid
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: (604)..(604)
<223> OTHER INFORMATION: Xaa can be any naturally occurring amino acid
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: (606)..(606)
<223> OTHER INFORMATION: Xaa can be any naturally occurring amino acid
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: (718)..(718)
<223> OTHER INFORMATION: Xaa can be any naturally occurring amino acid

```

```

<400> SEQUENCE: 45

```

```

Met Pro Thr Gly Arg Phe Glu Thr Met Arg Glu Trp Val His Asp Ala
1          5          10          15
Ile Ser Ala Gln Arg Asn Glu Leu Ser Leu Phe Ser Arg Tyr Val
20         25         30
Ala Gln Gly Lys Gly Ile Leu Xaa Ser His Gln Leu Ile Asp Glu Phe
35         40         45
Leu Lys Thr Val Lys Val Asp Gly Thr Thr Glu Asp Leu Lys Asn Arg
50         55         60
Pro Phe Met Lys Val Leu Gln Ser Ala Glu Glu Ala Ile Val Leu Pro
65         70         75         80
Pro Phe Val Ala Leu Ala Ile Arg Pro Arg Pro Gly Val Arg Glu Tyr
85         90         95
Val Arg Val Asn Val Tyr Glu Leu Ser Val Asp His Leu Thr Val Ser
100        105        110
Glu Tyr Leu Arg Phe Lys Glu Glu Leu Val Asn Gly His Ala Asn Gly
115        120        125

```

-continued

---

Asn Tyr Leu Leu Glu Leu Asp Phe Glu Pro Phe Asn Ala Thr Phe Pro  
 130 135 140

Arg Pro Thr Arg Ser Ser Ser Ile Gly Asn Gly Val Gln Phe Leu Asn  
 145 150 155 160

Arg His Leu Ser Ser Ile Met Phe Arg Asn Lys Asp Ser Leu Glu Pro  
 165 170 175

Leu Leu Glu Phe Leu Arg Thr His Lys His Asp Gly Arg Ala Met Met  
 180 185 190

Leu Asn Asp Arg Ile Gln Asn Ile Arg Thr Leu Gln Glu Ala Leu Ala  
 195 200 205

Arg Ala Glu Glu Phe Leu Ser Lys Leu Pro Leu Ala Thr Pro Tyr Ser  
 210 215 220

Glu Phe Glu Phe Xaa Leu Gln Gly Met Gly Phe Glu Arg Gly Trp Gly  
 225 230 235 240

Asp Thr Xaa Gln Lys Val Ser Glu Met Val His Leu Leu Leu Asp Ile  
 245 250 255

Leu Gln Ala Pro Asp Pro Ser Val Leu Glu Thr Phe Leu Gly Arg Ile  
 260 265 270

Pro Met Val Phe Asn Val Val Ile Leu Ser Pro His Gly Tyr Phe Gly  
 275 280 285

Gln Ala Asn Val Leu Gly Leu Pro Asp Thr Gly Gly Gln Val Val Tyr  
 290 295 300

Ile Leu Asp Gln Val Arg Ala Leu Glu Ser Glu Met Leu Leu Arg Ile  
 305 310 315 320

Gln Lys Gln Gly Leu Asp Val Thr Pro Lys Ile Leu Ile Val Thr Arg  
 325 330 335

Leu Ile Pro Glu Ala Glu Gly Thr Thr Cys Asn Gln Arg Leu Glu Lys  
 340 345 350

Val Ser Gly Thr Glu His Xaa His Ile Leu Arg Ile Pro Phe Arg Thr  
 355 360 365

Glu Lys Gly Ile Leu Arg Lys Trp Ile Ser Arg Phe Asp Val Trp Pro  
 370 375 380

Tyr Leu Glu Thr Phe Ala Glu Asp Ala Ser Asn Glu Ile Ala Ala Glu  
 385 390 395 400

Leu Gln Gly Val Pro Asn Leu Ile Ile Gly Asn Tyr Ser Asp Gly Asn  
 405 410 415

Leu Val Ala Ser Leu Leu Ala Cys Lys Leu Gly Val Ile Gln Cys Asn  
 420 425 430

Ile Ala His Ala Leu Glu Lys Thr Lys Tyr Pro Glu Ser Asp Ile Tyr  
 435 440 445

Trp Arg Asn His Glu Asp Lys Tyr His Phe Ala Ser Gln Phe Thr Ala  
 450 455 460

Asp Leu Ile Ala Met Asn Asn Ala Asp Phe Ile Ile Thr Ser Thr Tyr  
 465 470 475 480

Gln Glu Ile Ala Gly Ser Lys Asn Lys Val Gly Gln Tyr Glu Ser His  
 485 490 495

Thr Ala Phe Thr Leu Pro Gly Leu Tyr Arg Val Val His Gly Ile Asn  
 500 505 510

Val Phe Asp Pro Lys Phe Asn Ile Val Ser Pro Gly Ala Asp Met Thr  
 515 520 525

Ile Tyr Phe Pro Tyr Ser Asp Lys Glu Arg Arg Leu Thr Ala Leu His  
 530 535 540

-continued

Glu Ser Ile Glu Glu Leu Leu Phe Ser Ser Glu Gln Asn Val Glu His  
 545 550 555 560  
 Val Gly Phe Leu Ser Asp Gln Xaa Lys Pro Ile Ile Phe Ser Met Ala  
 565 570 575  
 Arg Leu Asp Arg Val Lys Asn Leu Thr Gly Leu Val Glu Cys Tyr Ala  
 580 585 590  
 Xaa Asn Xaa Xaa Leu Arg Glu Val Ala Asn Leu Xaa Val Xaa Gly Gly  
 595 600 605  
 Tyr Val Asp Val Asn Gln Ser Arg Asp Arg Glu Glu Met Ala Glu Ile  
 610 615 620  
 Gln Lys Met His Ser Leu Ile Lys Gln Tyr Gly Leu His Gly Glu Phe  
 625 630 635 640  
 Arg Trp Ile Ala Ala Gln Met Asn Arg Ala Arg Asn Gly Glu Leu Tyr  
 645 650 655  
 Arg Tyr Ile Ala Asp Thr Lys Gly Val Phe Val Gln Pro Ala Phe Tyr  
 660 665 670  
 Glu Ala Phe Gly Leu Thr Val Val Glu Ser Met Thr Cys Gly Leu Pro  
 675 680 685  
 Thr Phe Ala Thr Cys His Gly Gly Pro Ala Glu Ile Ile Glu Asn Gly  
 690 695 700  
 Val Ser Gly Phe His Ile Asp Pro Tyr His Pro Glu Gln Xaa Ala Thr  
 705 710 715 720  
 Thr Leu Val Ser Phe Phe Glu Thr Cys Asn Ala Asp Pro Ser His Trp  
 725 730 735  
 Glu Lys Ile Ser Asp Gly Gly Leu Lys Arg Ile Tyr Glu Arg Tyr Thr  
 740 745 750  
 Trp Lys Lys Tyr Ser Glu Arg Leu Leu Thr Leu Ala Gly Val Tyr Ser  
 755 760 765  
 Phe Trp Lys His Val Ser Lys Leu Glu Arg Arg Glu Thr Arg Arg Tyr  
 770 775 780  
 Leu Glu Met Phe Tyr Ser Leu Lys Tyr Arg Asp Leu Ala Asn Ser Ile  
 785 790 795 800  
 Pro Leu Ala Thr Asp Glu His  
 805

<210> SEQ ID NO 46  
 <211> LENGTH: 26  
 <212> TYPE: DNA  
 <213> ORGANISM: Artificial sequence  
 <220> FEATURE:  
 <223> OTHER INFORMATION: Primer a

<400> SEQUENCE: 46

ccttgcaaaa cttaagatca aaagtc

26

<210> SEQ ID NO 47  
 <211> LENGTH: 26  
 <212> TYPE: DNA  
 <213> ORGANISM: Artificial sequence  
 <220> FEATURE:  
 <223> OTHER INFORMATION: Primer b

<400> SEQUENCE: 47

ctatagatgg gatgaagctg ctctcg

26

<210> SEQ ID NO 48  
 <211> LENGTH: 24  
 <212> TYPE: DNA

-continued

<213> ORGANISM: Artificial sequence  
 <220> FEATURE:  
 <223> OTHER INFORMATION: Primer c

<400> SEQUENCE: 48

agagaggagc tcattgcggt gagc 24

<210> SEQ ID NO 49  
 <211> LENGTH: 24  
 <212> TYPE: DNA  
 <213> ORGANISM: Artificial sequence  
 <220> FEATURE:  
 <223> OTHER INFORMATION: Primer d

<400> SEQUENCE: 49

cccattcaag catcgtctcg aacc 24

<210> SEQ ID NO 50  
 <211> LENGTH: 2234  
 <212> TYPE: DNA  
 <213> ORGANISM: Brassica napus  
 <220> FEATURE:  
 <221> NAME/KEY: misc\_feature  
 <222> LOCATION: (859)..(859)  
 <223> OTHER INFORMATION: n is a, c, g, or t

<400> SEQUENCE: 50

ctatagatgg gatgaagctg ctctcgacaa atctgataaa actaagaag gttagtaatc 60  
 aatttttaca aaatcataga ttattttttt cattgaatta tttttatgct ataccaagaa 120  
 ttgtatttta gtatttgttt taactacata taatagaatt aactacatat aaattaacta 180  
 aacttaaaat aaaaatagat ttgtttcctg aaattatttt aagaatatat atgtatatat 240  
 ctaaaatctt agacttagat agatttttct atctatctat tttggttact taaaataaat 300  
 aaatttgat aaataattgt atagttatca aaaattaaaa ctaatttttt taaagttggt 360  
 gatataataa atactaaaga tttaacgatt aagtatttat ttaagtatag aattttggtt 420  
 tttttttaag tttagttatg aagttgtaa ttatattaaa acaaaacaat atttcgaaat 480  
 tttattatca tttcgaata tatttttttt agtgatgatg tatgaattat ttcataaatt 540  
 tgaaagttaa ctaaaaaata tatcaacatg aattgtaata tatgagttat taccttaacc 600  
 aaaattataa attaacatta aatataatta tatatgtcat atttagccat acaatgtgtc 660  
 atcaatatta atagtcatgt caatattaca taatgccaat attatgctac ttaaacccca 720  
 aatcccctaa ctcccgttaa gtgcccacaa tcataaatat acttattcga caaaaataaaa 780  
 aactttaaaa tatttactaa tccgaccatg cacaagcatc cattccctat tccattgcca 840  
 cgggataaca atgcaaccna ctctcaaaa aaagaaaaat tcaagctctt ttgcaaaaaa 900  
 aaataaaaata attttaacac ctaaaaattt ttgtttccaa acttctacag ggaacacaca 960  
 taaaagaaaa agaggagctc cactcggatc acgcaacaaa ccaaaagggtg tgtcatgact 1020  
 cctaagatat aatatttctt tattcaaaat cataccattt taaattatga atgtatttcg 1080  
 tagtccacca gatatgtaat ccaccagcgt tcaaaccaaa gttttatgat tgtaagttaa 1140  
 agtgaattat aataatata tcttcacggt atcttttcat aactaattga gttatcaaac 1200  
 ttgatcgcac atgtggcttt gatagggtg acttttatgg tatacaatc tttcaaccta 1260  
 aaaacattat tgttctcctaa tatcttcat tatgcttgac tgcaacaaaa tattttctca 1320  
 tctgtttctt tcctttaaac caatttatta tcatctattt cctgacattt taatccatcc 1380  
 acctatgtca aaaacttata gaaaatgtca acttccaaac aaaacataat tgaacttcgc 1440

-continued

---

```

aaataaatc ttaataatat taaaaaatgt tacttaatta tttcttcaac cccattttcc 1500
gcgcgtagcg cggacaaaga ctctagttaa atatagaagt ttccgattct catcgataa 1560
aacggtgact ttggcgggct ttcattgtga acaaattggt ttaacaaacc actgcctagt 1620
cgtttagtgt agaatcagcg catggaactc cgattggagc gtgactttca cgtgcccggag 1680
gcccaccacc acagcggggcg ttacgctcta agaatctcgc ccacggtttt ctteactctgc 1740
cccccgccaa gtgtcttctc cgttcgcacc ttctcaccaa gttacaggaa ccctaaaaat 1800
ggcctttctt cagccccggc tataatacac acatgatcct atagtggggt cttccacaag 1860
ttacatctcc ttctggattg tacatttcaa gtgtttgtgt tttttctgcc tctgagagaa 1920
aatcatgccg acgggtaggt tcgagacgat gcgtgaatgg gttcacgacg ccatctctgc 1980
tcaacgcaat gagctcctct ctctcttttc caggatctc tctctctctt actgaatatg 2040
cgttacatat ataagttcag tacatgcatt gtcactttgt caactttcaa cagttgagag 2100
tagagcatgt taaaaaaaa agtttagttcg ttttacttgc atgtgtgttg tggttagtct 2160
caggaggagt aatgctttgg tttgctatgt ttagatacgt agctcagggg aaagggatac 2220
tgcagtccca ccag 2234

```

```

<210> SEQ ID NO 51
<211> LENGTH: 632
<212> TYPE: DNA
<213> ORGANISM: Brassica napus

```

```

<400> SEQUENCE: 51

```

```

ccaatttatt atcatctatt tctgacatt ttaatccatc cacctatgtc aaaaacttat 60
agaaaatgtc aacttccaaa caaacataa ttgaacttcg caaataaatt cttaataata 120
ttaaaaaatg ttacttaatt atttcttcaa cccatttttc cgcgcgtagc gcggacaaag 180
actctagtta aatatagaag tttccgattc tcatcgata aaacggtgac tttggcgggc 240
tttcatgtgt aacaaattgg ttaacaaac cactgcctag tcgttttagtg tagaatcagc 300
gcatggaact ccgattggag cgtgactttc acgtrccgga ggccccaccac cwcagcgggc 360
gttacgctct aagaatctcg cccacgggtt tcttcatctc cccccgcca agtgtctccc 420
tcgttcgcca cttctcatca tgttacaggg accataaaaa tggcgtatth cttcagcccc 480
gggtataaat acacacatga tctgtgtgtg ggttcttcca caagttacat ctctctctgg 540
tttttgatt gcaagtgttt gtattttttg cctccgagag aaaatcatgc cgaccggtag 600
gttcgagacg atgcgtgaat gggcctgaat tc 632

```

```

<210> SEQ ID NO 52
<211> LENGTH: 35
<212> TYPE: DNA
<213> ORGANISM: Artificial sequence
<220> FEATURE:
<223> OTHER INFORMATION: Primer SA188

```

```

<400> SEQUENCE: 52

```

```

ggcgcgcccc aatttattat catctatttc ctgac 35

```

```

<210> SEQ ID NO 53
<211> LENGTH: 34
<212> TYPE: DNA
<213> ORGANISM: Artificial sequence
<220> FEATURE:
<223> OTHER INFORMATION: Primer SA189

```

-continued

&lt;400&gt; SEQUENCE: 53

gcgccgcga tttctctcg gagcAAAA atac 34

&lt;210&gt; SEQ ID NO 54

&lt;211&gt; LENGTH: 36

&lt;212&gt; TYPE: DNA

&lt;213&gt; ORGANISM: Artificial sequence

&lt;220&gt; FEATURE:

&lt;223&gt; OTHER INFORMATION: Primer SA190

&lt;400&gt; SEQUENCE: 54

ggcgcgcct atagatggga tgaagctgct ctcgac 36

&lt;210&gt; SEQ ID NO 55

&lt;211&gt; LENGTH: 36

&lt;212&gt; TYPE: DNA

&lt;213&gt; ORGANISM: Artificial sequence

&lt;220&gt; FEATURE:

&lt;223&gt; OTHER INFORMATION: Primer SA191

&lt;400&gt; SEQUENCE: 55

gcgccgcga tttctctca gagcagaaa aaacac 36

&lt;210&gt; SEQ ID NO 56

&lt;211&gt; LENGTH: 4114

&lt;212&gt; TYPE: DNA

&lt;213&gt; ORGANISM: Artificial sequence

&lt;220&gt; FEATURE:

&lt;223&gt; OTHER INFORMATION: Plasmid BN SUS2 prom1/PCR blunt

&lt;400&gt; SEQUENCE: 56

cctgaattct gcagatatcc atcacactgg cggccgctcg agcatgcac tagagggccc 60

aatcgccct atagtgagtc gtattacaat tcaactggccg tcgttttaca acgtcgtgac 120

tgggaaaacc ctggcgttac ccaactaat cgccttgacg cacatcccc tttcgcagc 180

tggcgtaata gcgaagaggc ccgcaccgat cgccttccc aacagttgag cagcctatac 240

gtacggcagt ttaaggtta cacctataaa agagagagcc gttatcgtct gtttgggat 300

gtacagagtg atattattga cagccgggg cgacggatgg tgatccccct ggccagtgca 360

cgtctgctgt cagataaagt ctccctgtaa ctttaccgg tgggtgcatat cggggatgaa 420

agctggcgca tgatgaccac cgatatggcc agtgtgccc tctccgttat cggggaagaa 480

gtggctgac tcagccaccg cgaaaatgac atcaaaaacg ccattaacct gatggtctgg 540

ggaatataaa tgtcaggcat gagattatca aaaaggatct tcacctagat ccttttcacg 600

tagaaagcca gtccgcagaa acgggtgctga ccccgatga atgtcagcta ctgggctatc 660

tggacaaggg aaaacgcaag cgcaaagaga aagcaggtag cttgcagtgg gcttacatgg 720

cgatagctag actggggggt tttatggaca gcaagcgaac cggaattgcc agctggggcg 780

ccctctgcta aggttgggaa gccctgcaaa gtaaaactgga tggctttctt gccgccaagg 840

atctgatggc gcaggggatc aagctctgat caagagacag gatgaggatc gtttcgcatg 900

attgaacaag atggattgca cgcaggttct ccggccgctt ggggtggagag gctattcggc 960

tatgactggg cacaacagac aatcggtctc tctgatgccg ccgtggtccg gctgtcagcg 1020

cagggggccc cggttctttt tgtcaagacc gacctgtccg gtgccctgaa tgaactgcaa 1080

gacgaggcag cgcggctatc gtggctggcc acgacggggc ttccttgccc agctgtgctc 1140

gacgttgca ctgaagcggg aaggactgg ctgctattgg gcgaagtgcc ggggcaggat 1200

ctcctgtcat ctcacctgac tctgcccag aaagtatcca tcatggctga tgcaatgccc 1260

-continued

---

cggtgcata	cgcttgatcc	ggctacctgc	ccattcgacc	accaagcgaa	acatcgcatc	1320
gagcgagcac	gtactcggat	ggaagccggt	cttgtcgatc	aggatgatct	ggacgaagag	1380
catcaggggc	tcgcccagc	cgaactgttc	gccaggtcca	aggcagcat	gcccagcggc	1440
gaggatctcg	tcgtgaccca	tggcgatgcc	tgcttgccga	atatcatggt	ggaaaatggc	1500
cgctttctg	gattcatoga	ctgtggccgg	ctgggtgtgg	cggaccgcta	tcaggacata	1560
gcgttggtc	cccgtgatat	tgctgaagag	cttgccggcg	aatgggctga	ccgcttctc	1620
gtgctttacg	gtatcggcg	tcccgatctg	cagcgcctcg	ccttctatcg	ccttcttgac	1680
gagttctct	gaattattaa	cgcttaaca	ttcctgatgc	ggtatctct	ccttacgcat	1740
ctgtcggta	ttcacaccg	catcaggtg	cacttttcgg	ggaaatgtgc	gcggaacccc	1800
tatttgttta	ttttctaaa	tacattcaaa	tatgtatccg	ctcatgagac	aataaccctg	1860
ataaatgctt	caataatagc	acgtgaggag	ggccaccatg	gccaaagtga	ccagtgcctg	1920
tccggtgctc	accgcgcgcg	acgtcgcgg	agcggctcag	ttctggaccg	accggtcgg	1980
gttctcccgg	gacttcgtgg	aggacgactt	cgccgggtg	gtccgggacg	acgtgaccct	2040
gttcatcagc	gcggtccagg	accaggtggt	gccggacaac	accctggcct	gggtgtgggt	2100
gcgcggcctg	gacgagctgt	acgccgagtg	gtcggaggtc	gtgtccacga	acttccggga	2160
cgctcccggg	ccggccatga	ccgagatcgg	cgagcagccg	tgggggcggg	agttcgcct	2220
gcgcgaccgg	gccggcaact	gcgtgcactt	cgtggccgag	gagcaggact	gacacgtgct	2280
aaaacttcat	ttttaattta	aaaggatcta	ggtgaagatc	ctttttgata	atctcatgac	2340
caaaatccct	taacgtgagt	tttcgtcca	ctgagcgtca	gaccccgtag	aaaagatcaa	2400
aggatctct	tgagatcctt	ttttctgcg	cgtaatctgc	tgcttgcaaa	caaaaaaac	2460
accgtacca	gcggtggttt	gtttgccgga	tcaagageta	ccaactcttt	ttccgaaggt	2520
aactggcttc	agcagagcgc	agataccaaa	tactgttctt	ctagtgtagc	cgtagttagg	2580
ccaccacttc	aagaactctg	tagcacggcc	tacatacctc	gctctgctaa	tcctgttacc	2640
agtggctgct	gccagtgccg	ataagctgtg	tcttaccggg	ttggactcaa	gacgatagtt	2700
accggataag	gcgcagcgg	cgggctgaac	gggggggttcg	tgcacacagc	ccagcttgga	2760
gcgaacgacc	tacaccgaac	tgagatacct	acagcgtgag	ctatgagaaa	gcgccacgct	2820
tcccgaaggg	agaaaagcgg	acaggtatcc	ggtaaagcgg	agggtcggaa	caggagagcg	2880
cacgagggag	cttccagggg	gaaaagcctg	gtatctttat	agtcctgtcg	ggtttcgcca	2940
cctctgactt	gagcgtcgat	ttttgtgatg	ctcgtcaggg	ggcggagcc	tatggaaaa	3000
cgccagcaac	gcggcctttt	tacggttcct	ggccttttgc	tggccttttg	ctcacatggt	3060
ctttcctgcg	ttatcccctg	attctgtgga	taaccgtatt	accgctttg	agtgagctga	3120
taccgctcgc	cgacccgaa	cgaccgagcg	cagcagatca	gtgagcgagg	aagcggaaga	3180
gcgccaata	cgaaaaccgc	ctctccccgc	gcgttgcccg	attcattaat	gcagctggca	3240
cgacaggttt	cccactgga	aagcgggag	tgagcgcaac	gcaattaatg	tgagttagct	3300
cactcattag	gcaccccagg	ctttacactt	tatgcttccg	gctcgtatgt	tgtgtggaat	3360
tgtgagcgg	taacaatttc	acacagggaa	cagctatgac	catgattacg	ccaagctatt	3420
taggtgacgc	gtagaatac	tcaagctatg	catcaagctt	ggtaccgagc	tcggatccac	3480
tagtaacggc	cgccagtggt	ctggaattca	ggggcgcgcc	ccaatttatt	atcatctatt	3540
tcctgacatt	ttaatccatc	cacctatgtc	aaaaacttat	agaaaatgtc	aacttccaaa	3600

-continued

---

```

caaaacataa ttgaacttcg caaataaatt cttaataata ttaaaaaatg ttacttaatt 3660
atttcttcaa ccccatcttc cgcgcgtagc gcggacaaag actctagtta aatatagaag 3720
tttccgattc tcatcgata aaacggtgac ttggcgggc tttcatgtgt aacaaattgg 3780
ttaaacaac cactgcctag tcgtttagt tagaatcagc gcatggaact ccgattggag 3840
cgtgacttcc acgtgccgga ggcccaccac cacagcgggc gttacgctct aagaatctcg 3900
ccccaggttt tcttcatctc ccccccgcca agtgtctccc tcgttcgcca cttctcatca 3960
tgttacaggg accataaaaa tggcgtatct cttcagcccc gggataaat acacacatga 4020
tcctgtgggt ggttcttcca caagttacat ctccctctgg tttttgtatt gcaagtgttt 4080
gtattttttg cctccgagag aaaatcgcg cgcg 4114

```

```

<210> SEQ ID NO 57
<211> LENGTH: 5452
<212> TYPE: DNA
<213> ORGANISM: Artificial sequence
<220> FEATURE:
<223> OTHER INFORMATION: Plasmid BN SUS2 prom2/PCR blunt
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: (4379)..(4379)
<223> OTHER INFORMATION: n is a, c, g, or t

```

```

<400> SEQUENCE: 57

```

```

cctgaattct gcagatatcc atcacactgg cggccgctcg agcatgcac tagagggccc 60
aatcgccct atagtgagtc gtattacaat tcaactggcc tcgttttaca acgtcgtgac 120
tgggaaaacc ctggcgttac ccaactaat cgccttgac cacatcccc ttccgccc 180
tggcgtaata gcgaagaggc ccgcaccgat cgccttccc aacagttgcg cagcctatac 240
gtacggcagt ttaaggttta cacctataaa agagagagcc gttatcgtct gtttgtggat 300
gtacagagtg atattattga cacgccgggg cgacggatgg tgatccccct ggccagtgca 360
cgtctgctgt cagataaagt ctcccgtgaa ctttaccggc tgggtgcatat cggggatgaa 420
agctggcgca tgatgaccac cgatatggcc agtgtgccc tctccgttat cggggaagaa 480
gtggctgac tcagccaccg cgaaaatgac atcaaaaacg ccattaacct gatggtctgg 540
ggaatataaa tgtcaggcat gagattatca aaaaggatct tcacctagat ccttttcacg 600
tagaaagcca gtcgcgagaa acgggtgctga ccccgatga atgtcagcta ctgggctatc 660
tggacaaggg aaaacgcaag cgcaaagaga aagcaggtag cttgcagtgg gcttacatgg 720
cgatagctag actggggcgt tttatggaca gcaagcgaac cggaattgcc agctggggcg 780
ccctctggta aggttgggaa gccctgcaaa gtaaaactgga tggctttctt gccgccaagg 840
atctgatggc gcaggggatc aagctctgat caagagacag gatgaggatc gtttcgcatg 900
attgaacaag atggattgca cgcaggttct ccggccgctt gggtggagag gctattcggc 960
tatgactggg cacaacagac aatcggctgc tctgatgcc ccggtgttccg gctgtcagcg 1020
cagggggccc cggttctttt tgtcaagacc gacctgtccg gtgccctgaa tgaactgcaa 1080
gacgaggcag cgcggctatc gtggctggcc acgacggggc ttccttgcgc agctgtgctc 1140
gacgttgcca ctgaagcggg aagggactgg ctgctattgg gcgaagtgcc ggggcaggat 1200
ctcctgtcat ctcacctgac tctgcccag aaagtatcca tcatggctga tgcaatgcgg 1260
cggtgcata cgcttgatcc ggctacctgc ccattcgacc accaagcgaa acatcgcac 1320
gagcgagcac gtactcggat ggaagccggt cttgtcgatc aggatgatct ggacgaagag 1380
catcaggggc tcgcgccagc cgaactgttc gccaggtcca aggcgagcat gcccgacggc 1440

```

-continued

---

gaggatctcg	tcgtgaccca	tggcgatgcc	tgcttgccga	atatcatggt	ggaaaatggc	1500
cgcttttctg	gattcatoga	ctgtggcccg	ctgggtgtgg	cggaccgcta	tcaggacata	1560
gcgttggtta	cccgtgatat	tgctgaagag	cttggcgcg	aatgggctga	ccgcttctc	1620
gtgctttacg	gatacgccgc	tcccgattcg	cagcgcatcg	ccttctatcg	ccttcttgac	1680
gagttcttct	gaattattaa	cgcttacaat	ttcctgatgc	ggtattttct	ccttacgcat	1740
ctgtgcggta	ttcacaccg	catcaggtgg	cacttttcgg	ggaaatgtgc	gcggaacccc	1800
tatttgttta	tttttctaaa	tacattcaaa	tatgtatccg	ctcatgagac	aataaccctg	1860
ataaatgctt	caataatagc	acgtgaggag	ggccaccatg	gccaaagtga	ccagtgccgt	1920
tccgtgtctc	accgcgcgcg	acgtcgcccg	agcggtcgag	ttctggaccg	accggctcgg	1980
gttctcccgg	gacttcgtgg	aggacgactt	cgccgggtg	gtccgggacg	acgtgaccct	2040
gttcatcagc	gcggtccagg	accaggtggt	gccggacaac	accctggcct	gggtgtgggt	2100
gcgcggcctg	gacgagctgt	acgccgagt	gtcggaggtc	gtgtccacga	acttccggga	2160
cgctcccggg	ccggccatga	ccgagatcgg	cgagcagccg	tgggggcggg	agttcgccct	2220
gcgcgaccog	gccggcaact	gcgtgcactt	cgtggccgag	gagcaggact	gacacgtgct	2280
aaaacttcat	ttttaattta	aaaggatcta	ggtgaagatc	ctttttgata	atctcatgac	2340
caaaatccct	taacgtgagt	tttcgttcca	ctgagcgtca	gaccccgtag	aaaagatcaa	2400
aggatcttct	tgagatcctt	ttttctgcg	cgtaatctgc	tgcttgcaaa	caaaaaaac	2460
accgtacca	gcggtggttt	gtttgccgga	tcaagageta	ccaactcttt	ttccgaaggt	2520
aactggcttc	agcagagcgc	agataccaaa	tactgttctt	ctagtgtagc	cgtagttagg	2580
ccaccacttc	aagaactctg	tagcacccgc	tacatacctc	gctctgctaa	tcctgttacc	2640
agtggctgct	gccagtgccg	ataagtcgtg	tcttaccggg	ttggactcaa	gacgatagtt	2700
accggataag	gcgcagcgg	cgggctgaac	gggggggtcg	tgcacacagc	ccagcttggg	2760
gcgaacgacc	tacaccgaac	tgagatacct	acagcgtgag	ctatgagaaa	gcgccacgct	2820
tcccgaaggg	agaaagcgg	acaggtatcc	ggtgaagcgc	agggtcggaa	caggagagcg	2880
cacgagggag	cttccagggg	gaaaagcctg	gtatctttat	agtcctgtcg	ggtttcgcca	2940
cctctgactt	gagcgtcgat	ttttgtgatg	ctcgtcaggg	ggcggagacc	tatggaaaaa	3000
cgccagcaac	gcggcctttt	tacggttcct	ggccttttgc	tggccttttg	ctcacatggt	3060
ctttcctgcg	ttatcccctg	attctgtgga	taaccgtatt	accgctttg	agtgagctga	3120
taccgctcgc	cgacccgaa	cgaccgagcg	cagcgagtca	gtgagcgagg	aagcgggaaga	3180
gcgcccaata	cgaaaaccgc	ctctccccgc	gcgttgcccg	attcattaat	gcagctggca	3240
cgacaggttt	cccagctgga	aagcgggcag	tgagcgcac	gcaattaatg	tgagttagct	3300
cactcattag	gcaccccagg	ctttacactt	tatgcttccg	gctcgtatgt	tgtgtggaat	3360
tgtgagcggg	taacaatttc	acacaggaaa	cagctatgac	catgattacg	ccaagctatt	3420
taggtgacgc	gttagaatac	tcaagctatg	catcaagctt	ggtaccgagc	tcggatccac	3480
tagtaacggc	cgccagtgtg	ctggaattca	ggggcgcgcc	ctatagatgg	gatgaagctg	3540
ctctcgacaa	atctgataaa	actaaagaag	gttagtaaat	aatttttaca	aaatcataga	3600
ttattttttt	cattgaatta	tttttatgct	ataccaagaa	ttgtatttta	gtatttgttt	3660
taactacata	taatagaatt	aactacatat	aaattaacta	aacttaaaat	aaaaatagat	3720
ttgtttcctg	aaattatttt	aagaatatat	atgtatatat	ctaaaaatctt	agacttagat	3780

-continued

---

```

agatttttct atctatctat tttggttact taaaataaat aaatttgat aaataattgt 3840
atagttatca aaaattaaaa ctaatttttt taaagttggt gatataaaa atactaaaga 3900
tttaacgatt aagtatttat ttaagtatag aattttggtt tttttttaag tttagttatg 3960
aagttggtta ttatattaaa acaaaacaat atttcgaaat tttattatca tattcgaata 4020
tatttttttt agtgatgatg tatgaattat tatcataatt tgaagttta ctaaaaaata 4080
tatcaacatg aattgtaata tatgagttat taccttaacc aaaattataa attaacatta 4140
aatataatta tatatgcoat atttagccat acaatgtgtc atcaatatta atagtcatgt 4200
caatattaca taatgccaat attatgctac ttaaacccta aatcccctaa ctcccgtaa 4260
gtagccaaat tcataaatat acttattcga caaaataaaa aactttaaaa tatttactaa 4320
tccgaccatg cacaagcadc cattccctat tccattgcca cgggataaca atgcaacna 4380
ctcctcaaaa aaagaaaaat tcaagctctt ttgcaaaaaa aaataaaata attttaaac 4440
ctaaaatttt ttgtttccaa acttctacag ggaacacaca taaaagaaaa agaggacgtc 4500
cactcggatc acgcaacaaa ccaaaagggt gtgcatgact cctaagatat aatatttct 4560
tattcaaaat cataccattt taaattatga atgtatttgc tagtocacca gatatgtaat 4620
ccaccagcgt tcaaaccaaa gttttatgat tgtaagtta agtgaattat aataatata 4680
tcttcacggt atcttttcoat aactaattga gttatcaaac ttgatcgac atgtggctt 4740
gatagggtg acttttatgg tatacaattc tttcaacct aaaaattat tgttcctcaa 4800
tatcttcat tatgcttgac tgaacaaaa tatttttctc tctgttttct tctttaaac 4860
caatttatta tcctctattt cctgacattt taatccatcc acctatgtca aaaacttata 4920
gaaaatgtca acttccaac aaaacataat tgaacttcgc aaataaattc ttaataatat 4980
taaaaaatgt tacttaatta tttctcaac cccattttcc gcgctagcg cggacaaaga 5040
ctctagttaa atatagaagt tccgattct catcgtataa aacggtgact ttggcgggct 5100
ttcatgtgta acaaatgggt ttaacaaacc actgcctagt cgtttagtgt agaatecagc 5160
catggaactc cgattggagc gtgactttca cgtgcgggag gccaccacc acagcgggcg 5220
ttacgctcta agaactcgc ccacggttt ttctatctgc ccccgccaa gtgtcttct 5280
cgttcgcac ttctcaccia gttacaggaa cctaaaaat ggcttttct cagccccggc 5340
tataatacac acatgatcct atagtgggtt ctccacaag ttacatctcc ttctggattg 5400
tacatttcaa gtgtttgtgt tttttctgcc tctgagagaa aatcgcggcc gc 5452

```

```

<210> SEQ ID NO 58
<211> LENGTH: 8227
<212> TYPE: DNA
<213> ORGANISM: Artificial sequence
<220> FEATURE:
<223> OTHER INFORMATION: Vector KS427
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: (6408)..(6408)
<223> OTHER INFORMATION: n is a, c, g, or t

```

```

<400> SEQUENCE: 58

```

```

ctagagggcc caattcgccc tatagtgagt cgtattacaa ttcactggcc gtcgttttac 60
aacgtcgtga ctgggaaaac cctggcggtta cccaacttaa tcgccttgca gcacatcccc 120
cttctgccag ctggcgtaat agcgaagagg cccgcaccga tcgccttcc caacagttgc 180
gcagcctata cgtacggcag ttaagggtt acacctataa aagagagagc cgttatcgtc 240
tgtttggtga tgtacagagt gatattattg acacgcccgg gcgacggatg gtgatcccc 300

```

-continued

---

tggccagtgc	acgtctgctg	tcagataaag	tctcccgtga	actttaccgg	gtggtgcata	360
tcggggatga	aagctggcgc	atgatgacca	ccgatatggc	cagtgtgccg	gtctccgtta	420
tcggggaaga	agtggctgat	ctcagccacc	gcgaaaatga	catcaaaaac	gccattaacc	480
tgatgttctg	gggaatataa	atgtcaggca	tgagattatc	aaaaaggatc	ttcacctaga	540
tccttttcac	gtagaaagcc	agtcgcgaga	aacgggtgctg	accccggatg	aatgtcagct	600
actgggctat	ctggacaagg	gaaaacgcaa	gcgcaaagag	aaagcaggta	gcttgcagtg	660
ggcttacatg	gcgatagcta	gactggggcg	ttttatggac	agcaagcgaa	ccggaattgc	720
cagctggggc	gccctctggt	aaggttggga	agccctgcaa	agtaaactgg	atggctttct	780
tgcccacaag	gatctgatgg	cgcaggggat	caagctctga	tcaagagaca	ggatgaggat	840
cgtttcgcac	gattgaaaca	gatggattgc	acgcaggttc	tccggccgct	tggtgggaga	900
ggctattcgg	ctatgactgg	gcacaacaga	caatcggtg	ctctgatgcc	gccgtgttcc	960
ggctgtcagc	gcagggggcg	ccggttcttt	ttgtcaagac	cgacctgtcc	ggtgccctga	1020
atgaactgca	agacgaggca	gcgcggctat	cgtggctggc	cacgacgggc	gttccttgcg	1080
cagctgtgct	cgacgttgtc	actgaagcgg	gaagggactg	gctgctattg	ggcgaagtgc	1140
cggggcagga	tctcctgtca	tctcaccttg	ctcctgccga	gaaagtatcc	atcatggctg	1200
atgcaatcgc	gcggctgcat	acgcttgatc	cggtacctg	cccattcgac	caccaagcga	1260
aacatcgcac	cgagcgagca	cgtactcgga	tggaaagccg	tcttgtcgat	caggatgatc	1320
tggacgaaga	gcatcagggg	ctcgcgccag	ccgaactgtt	cgccaggctc	aaggcgagca	1380
tgcccgcagc	cgaggatctc	gtcgtgacct	atggcgatgc	ctgcttgccg	aatatcatgg	1440
tggaaaatgg	ccgcttttct	ggattcatcg	actgtggccg	gctgggtgtg	gcggaccgct	1500
atcaggacac	agcgttggct	accctgata	ttgctgaaga	gcttggcggc	gaatgggctg	1560
accgcttctc	cgtgctttac	ggtatcgccg	ctcccgatcc	gcagcgcate	gccttctatc	1620
gccttcttga	cgagttcttc	tgaattatta	acgcttacia	tttctgatg	cggtattttc	1680
tccttacgca	tctgtgcggt	atttcacacc	gcatcaggtg	gcacttttcc	gggaaatgtg	1740
cgcggaaccc	ctatttgttt	atttttctaa	atacattcaa	atatgtatcc	gctcatgaga	1800
caataaccct	gataaatgct	tcaataatag	cacgtgagga	gggcccaccat	ggccaagtgt	1860
accagtgcgc	ttccgggtgct	caccgcgcgc	gacgtcgcgc	gagcggctga	gttctggacc	1920
gaccggctcg	ggttctcccg	ggacttcgtg	gaggacgact	tcgcgggtgt	ggtccgggac	1980
gacgtgacct	tgttcatcag	cgcgggtccag	gaccaggtgg	tgccggacaa	caccctggcc	2040
tgggtgtggg	tgcgcggcct	ggacgagctg	tacgccaggt	ggtcggaggt	cgtgtccacg	2100
aacttccggg	acgctccggg	gccggccatg	accgagatcg	gcgagcagcc	gtggggggcg	2160
gagttcgcgc	tgcgcgaccc	ggccggcaac	tgcgtgcact	tcgtggccga	ggagcaggac	2220
tgacacgtgc	taaaaactca	tttttaattt	aaaaggatct	aggtgaagat	cctttttgat	2280
aatctcatga	ccaaaatccc	ttaacgtgag	ttttcgttcc	actgagcgtc	agaccccgtc	2340
gaaaagatca	aaggatcttc	ttgagatcct	ttttttctgc	gcgtaactctg	ctgcttgcaa	2400
acaaaaaac	caccgctacc	agcgggtggt	tgtttgccgg	atcaagagct	accaactctt	2460
tttccgaagg	taactggctt	cagcagagcg	cagataccaa	atactgttct	tctagtgtag	2520
ccgtagttag	gccaccactt	caagaactct	gtagcaccgc	ctacatacct	cgctctgcta	2580
atcctgttac	cagtggctgc	tgccagtggc	gataagtcgt	gtcttaccgg	gttggactca	2640

-continued

---

agacgatagt	taccggataa	ggcgcagcgg	tcgggctgaa	cggggggttc	gtgcacacag	2700
cccagcttgg	agcgaacgac	ctacaccgaa	ctgagatacc	tacagcgtga	gctatgagaa	2760
agcgccacgc	ttcccgaagg	gagaaaggcg	gacaggtatc	cggtaaagcg	cagggctcga	2820
acaggagagc	gcacgagggg	gcttccaggg	ggaaaagcct	ggtatcttta	tagtctgtc	2880
gggtttcgcc	acctctgact	tgagcgtcga	tttttggat	gctcgtcagg	ggggcggagc	2940
ctatgaaaa	acgccagcaa	cgcggccttt	ttacggttcc	tggccttttg	ctggcctttt	3000
gctcacatgt	tctttctgc	gttatccct	gattctgtgg	ataaccgtat	taccgccttt	3060
gagtgagctg	ataccgctcg	ccgcagccga	acgaccgagc	gcagcagctc	agtgagcag	3120
gaagcggaa	agcgccea	acgcaaaccg	cctctccccg	cgcgttgccc	gattcattaa	3180
tgacgctggc	acgacaggtt	tcccgaactg	aaagcgggca	gtgagcgcga	cgcaattaat	3240
gtgagttagc	tcactcatta	ggcaccocag	gctttact	ttatgcttcc	ggctcgtatg	3300
ttgtgtggaa	ttgtgagcgg	ataacaattt	cacacaggaa	acagctatga	ccatgattac	3360
gccaagctat	ttaggtgacg	cgttagaata	ctcaagctat	gcatcaagct	tggtaccgag	3420
ctcggatcca	ctagtaacgg	ccgccagtg	gctggaattc	aggggcgcgc	cccaatttat	3480
tatcatctat	ttcctgacat	tttaatccat	ccacctatgt	caaaaactta	tagaaaaatg	3540
caacttcaa	acaaaacata	attgaacttc	gcaataaat	tcttaataat	attaaaaaat	3600
gttacttaat	tatttcttca	accccat	ccgcgcgtag	cgcggacaaa	gactctagtt	3660
aaatatagaa	gtttccgatt	ctcatcgtat	aaaacgggta	ctttggcggg	ctttcatgtg	3720
taacaaattg	gtttaacaaa	ccactgccta	gtcgtttagt	gtagaatcag	cgcatggaac	3780
tccgattgga	gcgtgacttt	cacgtgccgg	aggcccacca	ccacagcggg	cgttacgctc	3840
taagaatctc	gccacaggtt	ttcttcatct	ccccccgcc	aagtgtctcc	ctcgttcgcc	3900
acttctcatc	atgttacagg	gaccataaaa	atggcgtatt	tcttcagccc	cgggtataaa	3960
tacacacatg	atcctgtgg	gggttcttcc	acaagttaca	tctccttctg	gtttttgtat	4020
tgcaagtgtt	tgtatttttt	gctcccgaga	gaaaatcgcg	gccgcaagta	tgaactaaaa	4080
tgcatgtagg	tgtaagagct	catggagagc	atggaatatt	gtatccgacc	atgtaacagt	4140
ataataactg	agctccatct	cacttcttct	atgaataaac	aaaggatggt	atgatataat	4200
aacactctat	ctatgcacct	tattgttcta	tgataaattt	cctcttatta	ttataaatca	4260
tctgaatcgt	gacggcttat	ggaatgcttc	aaatagtaca	aaaacaaatg	tgtactataa	4320
gactttctaa	acaattctaa	ccttagcatt	gtgaacgaga	cataagtgtt	aagaagacat	4380
aacaattata	atggaagaag	ttgtctcca	tttatataat	atatattacc	cacttatgta	4440
ttatattagg	atgttaagga	gacataacaa	ttataaagag	agaagtttgt	atccatttat	4500
atattatata	ctaccat	atatattata	cttatccact	tatttaatgt	ctttataagg	4560
tttgatccat	gatatttcta	atattttagt	tgatatgtat	atgaaaggg	actatttgaa	4620
ctctcttact	ctgtataaag	gttgatcat	ccttaaagtg	ggtctattta	attttattgc	4680
ttcttacaga	taaaaaaaaa	attatgagtt	ggttgataa	aatattgaag	gatttaaaat	4740
aataataaat	aacatataat	atatgtatat	aaatttatta	taatataaca	tttatctata	4800
aaaaagtaaa	tattgtcata	aatctataca	atcgtttagc	cttgctggac	gaatctcaat	4860
tatttaaacg	agagtaaac	tatttgactt	tttggttatt	taacaaatta	ttatttaaca	4920
ctatatgaaa	ttttttttt	tatcagcaaa	gaataaaatt	aaattaagaa	ggacaatggt	4980
gtcccaatcc	ttatacaacc	aacttcaca	agaaagtcaa	gtcagagaca	acaaaaaac	5040

-continued

---

aagcaaagga aatTTTTTaa tttgagttgt cttgtttgct gcataattta tgcagtaaaa	5100
cactacacat aaccctTTta gcagtagagc aatggttgac cgtgtgctta gcttctTTta	5160
TTTTattttt ttatcagcaa agaataaata aaataaaatg agacacttca gggatgtttc	5220
aacaagcttg gatcctcgaa gagaagggtt aataaacacac tttTTTaaCa tttTTTaaCa	5280
aaatTTtagt tattTaaaaa tttattaaaa aatTTaaaa aagaagagga actcTTTaaa	5340
TaaatcTaaC ttacaaaatt tatgattttt aataagtttt caccaataaa aaatgTcata	5400
aaaatagtTt aaaaagtata ttatcaatat tctcTTtatg ataaataaaa agaaaaaaa	5460
aataaaagtT aagtGaaaat gagattgaag tgactTTagg tgtgtataaa tatatCaacc	5520
ccGcCaCaCa tttattTaat ccaaatatat tgaagtatat tattocatag cctttattTta	5580
tttatatatt tattatataa aagcTTtatt tgttctaggT tgttcatgaa atattTTTTt	5640
ggTTTTatct ccgttGtaaG aaaatcatgt gctttgtgtc gccactcact attgcagctT	5700
tttcatgcat tggtcagatt gacggttgat tgtattTTTg tttTTtatgg ttttGtgTta	5760
tgactTaaGt cttcatctct ttatctcttc atcaggTTtg atggtTacct aatatggTcc	5820
atgggtacat gcattgTtaa attaggtggc caactTTgtt gtgaacgata gaattTTTTt	5880
tatattaaGt aaactatttt tatattatga aataataata aaaaaatat tttatcatta	5940
TtaCaCaCaat catattagtt aatttGttaa ctctataata aaGaaatac tGtaacattc	6000
acattacatg gtaacatctt tccacccttT catttGtttT ttgtttgatg actTTTTttc	6060
ttgtTTaaat ttatttccct tctTTTaaat ttGgaataca ttatcatcat atataaacta	6120
aaatactaaa aacaggatta cacaatgat aaataataac acaaatattt ataaatctag	6180
ctgcaataata tttaaactag ctatatogat attgtaaaat aaaactagct gcattgatac	6240
Tgataaaaaa atatcatgtg ctttctggac tgatgatgca gtatactttt gacattgcct	6300
ttattttatt tttcagaaaa gctttcttag ttctgggttc ttcattattt gtttcccatc	6360
tccattgtga attgaatcat ttgcttctgt tcacaatac aatttagnta ggtacatgca	6420
ttggtcagat tcacggTTta ttatgTcatg actTaaGttc atggtagtac attacctgcc	6480
acgcatgcat tatattggtt agatttgata ggcaaatTTg gttgtcaaca atataaatat	6540
aaataatgtt tttatattac gaaataacag tgatCaCaCaac aaacagtttt atctttatta	6600
acaagatttt gtttttGttt gatgacgttt tttaatgttt acgctttccc ccttctTTTg	6660
aatTTtagaac actttatcat cataaaatca aatactaaaa aaattacata tttcataaat	6720
aataacacaa atattTTTaa aaaatctgaa ataataatga acaatattac atattatcac	6780
Gaaaattcat taataaaaat attatataaa taaaatgtaa tagtagttat atgtaggaaa	6840
aaagtactgc acgcataata tatacaaaaa gattaaaatg aactattata aataataaca	6900
ctaaattaat ggtgaatcat atCaCaCaataa tGaaaaagta aataaaattt gtaattaact	6960
tctatatgta ttacacacac aaataataaa taatagTaaa aaaaattatg ataaatattt	7020
accatctcat aagatattta aaataatgat aaaaatatag attattTTTT atgcaactag	7080
ctagcCaCaCa agagaacacg ggtatatata aaaagagtac cttTaaattc tactgtactt	7140
cctttattcc tgacgTTTTt atatcaagtG gacatacgtg aagattTTaa ttatcagtct	7200
aaatattTca ttagcactta atacttttct gttttattcc tatcctataa gtagtcccga	7260
ttctcccaac attgcttatt cacacaacta actaagaaag tcttccatag ccccccaagc	7320
ggcccatggc ctctccgag gacgtcatca aggagtTcat gcgcttcaag gtgcgcatgg	7380

-continued

---

```

agggtcccg  gaacggccac  gagttcgaga  tcgagggcga  gggcgagggc  cgcccctacg  7440
agggcaccca  gaccgccaag  ctgaaggtga  ccaagggcgg  ccccctgccc  ttcgcctggg  7500
acatcctgtc  cccccagttc  cagtacggct  ccaaggtgta  cgtgaagcac  cccgcccaca  7560
tccccgacta  caagaagctg  tccttccccg  agggcttcaa  gtgggagcgc  gtgatgaact  7620
tcgaggacgg  cggcggtggt  accgtgaccc  aggactcctc  cctgcaggac  ggctccttca  7680
tctacaaggt  gaagttcctc  ggcgtgaact  tcccctccga  cggccccgta  atgcagaaga  7740
agactatggg  ctgggagggc  tccaccgagc  gcctgtaccc  ccgcgacggc  gtgctgaagg  7800
gcgagatcca  caaggccctg  aagctgaagg  acggcggcca  ctacctggtg  gagttcaagt  7860
ccatctacat  ggccaagaag  cccgtgcagc  tgcccggcta  ctactacgtg  gactccaagc  7920
tggacatcac  ctcccacaac  gaggactaca  ccatcgtgga  gcagtacgag  cgcgcccagg  7980
gccgccacca  cctgttctcg  tagcggcccg  ccgcgacaca  agtgtgagag  tactaaataa  8040
atgctttggt  tgtacgaaat  cattacacta  aataaaataa  tcaaagetta  tatatgcctt  8100
ccgctaaggc  cgaatgcaaa  gaaattggtt  ctttctcgtt  atcttttgc  acttttacta  8160
gtacgtatta  attactactt  aatcatcttt  gtttacggct  cattatatcc  gtcgacggcg  8220
cgccgct  8227

```

```

<210> SEQ ID NO 59
<211> LENGTH: 5704
<212> TYPE: DNA
<213> ORGANISM: Artificial sequence
<220> FEATURE:
<223> OTHER INFORMATION: Vector KS130

```

```

<400> SEQUENCE: 59

```

```

ggccgcaagt  atgaactaaa  atgcatgtag  gtgtaagagc  tcatggagag  catggaatat  60
tgtatccgac  catgtaacag  tataataact  gagctccatc  tcacttcttc  tatgaataaa  120
caaaggatgt  tatgatatat  taacactcta  tctatgcacc  ttattgttct  atgataaatt  180
tcctcttatt  attataaatc  atctgaatcg  tgacggctta  tggaatgctt  caaatagtac  240
aaaaacaaat  gtgtactata  agactttcta  aacaattcta  accttagcat  tgtgaacgag  300
acataagtgt  taagaagaca  taacaattat  aatggaagaa  gtttgtctcc  atttatatat  360
tatatattac  ccacttatgt  attatattag  gatgttaagg  agacataaca  attataaaga  420
gagaagtttg  tatccattta  tatattatat  actaccatt  tatatattat  acttatccac  480
ttatttaatg  tctttataag  gtttgatcca  tgatatttct  aatattttag  ttgatatgta  540
tatgaaaggg  tactatttga  actctcttac  tctgtataaa  ggttgatca  tccttaaagt  600
gggtctattt  aattttatgt  cttcttacag  ataaaaaaaa  aattatgagt  tggtttgata  660
aaatattgaa  ggatttaaaa  taataataaa  taacataata  tatatgtata  taaatttatt  720
ataatataac  atttatctat  aaaaaagtaa  atattgtcat  aaatctatac  aatcgtttag  780
ccttgctgga  cgaatctcaa  ttatttaaac  gagagtaaac  atatttgact  ttttggttat  840
ttaacaaatt  attatttaac  actatatgaa  attttttttt  ttatcagcaa  agaataaaat  900
taaattaaga  aggacaatgg  tgtcccaatc  cttatacaac  caacttcac  aagaaagtca  960
agtcagagac  aacaaaaaaaa  caagcaagg  aaatttttta  atttgagttg  tcttgtttgc  1020
tgcataattt  atgcagtaaa  aactacaca  taaccctttt  agcagtagag  caatggttga  1080
ccgtgtgctt  agcttctttt  atttatttt  tttatcagca  aagaataaat  aaaataaaat  1140
gagacacttc  aggatgttt  caacaagctt  ggatccgtcg  acggcgcgcc  cgatcatccg  1200

```

-continued

---

gatatagttc	ctcctttcag	caaaaaaccc	ctcaagacc	gttttagaggc	cccaaggggt	1260
tatgctagtt	attgctcagc	ggtggcagca	gccaaactcag	cttcctttcg	ggctttgtta	1320
gcagccggat	cgatccaagc	tgtacctcac	tattcctttg	ccctcggacg	agtgtgggg	1380
cgteggtttc	cactatcggc	gagtacttct	acacagccat	cggtccagac	ggcgcgctt	1440
ctgcgggcca	tttgtgtacg	cccgcagtc	ccggctccgg	atcggacgat	tgcgtcgc	1500
cgaccctgcg	cccaagctgc	atcatcgaaa	ttgccgtcaa	ccaagctctg	atagagttgg	1560
tcaagaccaa	tgcggagcat	atacgcggcg	agccgcggcg	atcctgcaag	ctcgggatgc	1620
ctccgctcga	agtagcgggt	ctgctgctcc	atacaagcca	accacggcct	ccagaagaag	1680
atgttgccga	cctcgtattg	ggaatccccg	aacatcgcc	cgctccagtc	aatgaccgct	1740
gttatgcggc	cattgtccgt	caggacattg	ttggagccga	aatccgcgtg	cacgaggtgc	1800
cggacttcgg	ggcagtcctc	ggcccaaagc	atcagctcat	cgagagcctg	cgcgacggac	1860
gcactgacgg	tgctgtccat	cacagtttgc	cagtgataca	catggggatc	agcaatcgcg	1920
catatgaaat	cacgccatgt	agtgtattga	ccgattcctt	gccgtccgaa	tgggcccgaac	1980
ccgctcgtct	ggctaagatc	ggccgcagcg	atcgcaccca	tagcctccgc	gaccggctgc	2040
agaacagcgg	gcagttcggg	ttcaggcagg	tcttgcaacg	tgacaccctg	tgacagcggg	2100
gagatgcaat	aggtcaggct	ctcgtgcaat	tccccaatgt	caagcacttc	cggaatcggg	2160
agcgcggccg	atgcaaagtg	ccgataaaca	taacgatctt	tgtagaaacc	atcggcgcag	2220
ctatttaacc	gcaggacata	tccacgccct	cctacatcga	agctgaaagc	acgagattct	2280
tcgccctcgg	agagctgcat	caggctggag	acgctgtcga	acttttcgat	cagaaacttc	2340
tcgacagacg	tcgcggtgag	ttcaggcttt	tccatgggta	tatctccttc	ttaaagttaa	2400
acaaaattat	ttctagaggg	aaaccgttgt	ggtctcccta	tagtgagtcg	tattaatttc	2460
gogggatcga	gatctgatca	acctgcatta	atgaaatcggc	caacgcgcgg	ggagagggcg	2520
tttgcgtatt	gggcgctott	ccgcttctc	gctcactgac	tcgctgcgct	cggtcgttcg	2580
gctgcggcca	gcggtatcag	ctcactcaaa	ggcggtaata	cggttatcca	cagaatcagg	2640
ggataaacgca	ggaaaagaaca	tgtgagcaaa	aggccagcaa	aaggccagga	accgtaaaaa	2700
ggccgcgctg	ctggcgtttt	tccatagggc	ccgccccct	gacgagcctc	acaaaaatcg	2760
acgctcaagt	cagaggtggc	gaaaccggac	aggactataa	agataaccag	cgtttcccc	2820
tggaagctcc	ctcgtgcgct	ctcctgttcc	gacctgccc	cttaccggat	acctgtccgc	2880
ctttctccct	tcgggaaagc	tggcgcttcc	tcaatgctca	cgctgtaggt	atctcagttc	2940
ggtgtaggtc	gttcgctcca	agctgggctg	tgtgcacgaa	cccccgcttc	agcccgaccg	3000
ctgcgcctta	tccggtaact	atcgtcttga	gtccaaccgg	gtaagacacg	acttatcgcc	3060
actggcagca	gccactggta	acaggattag	cagagcgagg	tatgtagggc	gtgctacaga	3120
gttcttgaag	tggtggccta	actacggcta	cactagaagg	acagtatttg	gtatctgcgc	3180
tctgctgaag	ccagttacct	tcggaaaaag	agttggtagc	tcttgatccg	gcaaacaaac	3240
caccgctggg	agcggtggtt	ttttgtttg	caagcagcag	attacgcgca	gaaaaaaagg	3300
atctcaagaa	gatcctttga	tctttctac	ggggtctgac	gctcagtgga	acgaaaaactc	3360
acgttaaggg	atthtggta	tgacattaac	ctataaaaat	aggcgtatca	cgaggccctt	3420
tcgtctcggc	cgtttcgggtg	atgacgggta	aaacctctga	cacatgcagc	tcccggagac	3480
ggtcacagct	tgtctgtaag	cggatgccgg	gagcagacaa	gcccgtcagg	gcgcgtcagc	3540

-continued

---

```

gggtgttggc ggggtgtcggg gctggcttaa ctatgcgcca tcagagcaga ttgtactgag 3600
agtgcaccaa atggacatat tgtcgttaga acgcggttac aattaatata taaccttatg 3660
tatcatacac atacgattta ggtgacacta tagaacggcg cgccaagctt ttgatccatg 3720
cccttcattt gccgcttatt aattaatttg gtaacagtcg gtactaatca gttacttatc 3780
cttcccccat cataattaat cttggtagt cgaatgcca caacactgac tagtctcttg 3840
gatcataaga aaaagccaag gaacaaaaga agacaaaaca caatgagagt atcctttgca 3900
tagcaatgtc taagttcata aaattcaaac aaaaacgcaa tcacacacag tggacatcac 3960
ttatccacta gctgatcagg atcgccgct caagaaaaaa aaactggacc ccaaaagcca 4020
tgcaacaaca cacgtactca caaaggtgtc aatcgagcag cccaaaacat tcaccaactc 4080
aaccatcat gagccctcac atttgttgt tctaaccaca cctcaactc gtattctctt 4140
ccgccacctc atttttgttt atttcaacac ccgtcaaaact gcatgccacc ccgtggccaa 4200
atgtccatgc atgttaacaa gacctatgac tataaatagc tgcaatctcg gccagggtt 4260
tcatcatcaa gaaccagttc aatattctag tacaccgat taaagaattt aagatatact 4320
gcgcccgcat ggctgctgct cccagtggtg ggacgtttac tcgggcccag gttttgaatg 4380
ccgaggctct gaatgagggc aagaaggatg ccgaggcacc cttcttgatg atcatcgaca 4440
acaaggtgta cgatgtccgc gagttcgtcc ctgatcatcc cgggtggaagt gtgattctca 4500
cgcacgttgg caaggacggc actgacgtct ttgacacttt tcaccccag gctgcttggg 4560
agactcttgc caacttttac gttggtgata ttgacgagag cgaccgcat atcaagaatg 4620
atgactttgc ggccgaggtc cgcaagctgc gtacctgtt ccagtctctt ggttactacg 4680
attcttccaa ggcatactac gccttcaagg tctcgttcaa cctctgcatc tggggtttgt 4740
cgacggteat tgtggccaag tggggccaga cctcgacct cgccaacgtg ctctcgctg 4800
cgcttttggg tctgttctgg cagcagtgcg gatggttggc tcacgacttt ttgcatcacc 4860
aggctctcca ggaccgttcc tggggtgatc ttttcggcgc cttcttggga ggtgtctgcc 4920
agggtctctc gtctctgtgg tggaggaca agcacaacac tcaccacgcc gcccccaacg 4980
tccacggcga ggatcccagc attgacacct accctctgtt gacctggagt gagcatgctg 5040
tggagatggt ctccgatgct ccagatgagg agctgacctg catgtggtcg cgtttcatgg 5100
tcttgaacca gacctggttt tacttcccca ttctctctgt tgcccgtctc tctggtgcc 5160
tccagtccat tctctttgtg ctgcctaacg gtcaggccca caagccctcg ggcgcgctg 5220
tgcccatctc gttggtcgag cagctgtcgc ttgcatgca ctggacctgg tacctcgcca 5280
ccatgttctc gttcatcaag gatcccgta acatgctggt gtactttttg gtgtcgcagg 5340
cgggtgtcgg aaacttgttg gcgatcgtgt tctcgtcaa ccacaacggt atgcctgtga 5400
tctcgaagga ggaggcggtc gatatggatt tcttcacgaa gcagatcacc acgggtcgtg 5460
atgtccaccc gggctatatt gccaaactgt tcacgggtgg attgaactat cagatcgagc 5520
accacttgtt cccttcgatg cctcgccaca acttttcaaa gatccagcct gctgtcgaga 5580
ccctgtgcaa aaagtacaat gtccgatacc acaccaccgg tatgatcgag ggaactgcag 5640
aggctcttag ccgtctgaac gaggtctcca aggtctcctc caagatgggt aaggcgcagt 5700
aagc 5704

```

&lt;210&gt; SEQ ID NO 60

&lt;211&gt; LENGTH: 9609

&lt;212&gt; TYPE: DNA

&lt;213&gt; ORGANISM: Artificial sequence

-continued

---

```

<220> FEATURE:
<223> OTHER INFORMATION: Vector KS432
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: (2349)..(2349)
<223> OTHER INFORMATION: n is a, c, g, or t

<400> SEQUENCE: 60
ggccgcaagt atgaactaaa atgcatgtag gtgtaagagc tcatggagag catggaatat    60
tgtatccgac catgtaacag tataataact gagctccatc tcaacttcttc tatgaataaa    120
caaaggatgt tatgatatat taacctcta tctatgcacc ttattgttct atgataaatt    180
tcctcttatt attataaatc atctgaatcg tgacggctta tggaaatgctt caaatagtac    240
aaaaacaat gtgtactata agactttcta aacaattcta accttagcat tgtgaacgag    300
acataagtgt taagaagaca taacaattat aatggaagaa gtttgcctcc atttatatat    360
tatatattac ccacttatgt attatattag gatgtaagg agacataaca attataaaga    420
gagaagtttg tatccattta tatattatat actaccatt tatatattat acttatccac    480
ttatttaatg tctttataag gtttgatcca tgatatttct aatattttag ttgatatgta    540
tatgaaaggg tactatttga actctcttac tctgtataaa ggttgatca tccttaaagt    600
gggtctattt aattttattg cttcttacag ataaaaaaaa aattatgagt tggtttgata    660
aaatattgaa ggatttaaaa taataataaa taacatataa tatatgtata taaatttatt    720
ataatataac atttatctat aaaaaagtaa atattgtcat aaatctatac aatcgtttag    780
ccttgctgga cgaatctcaa ttatttaaac gagagtaaac atatttgact ttttggttat    840
ttaacaaatt attatttaac actatatgaa attttttttt ttatcagcaa agaataaaat    900
taaattaaga aggacaatgg tgtccaatc cttatacaac caactccac aagaaagtca    960
agtcagagac aacaaaaaaaa caagcaagg aaatttttta atttgagttg tcttgtttgc   1020
tgcataatth atgcagtaaa acactacaca taaccctttt agcagtagag caatggttga   1080
ccgtgtgctt agcttctttt attttatttt tttatcagca aagaataaat aaaataaaat   1140
gagacacttc agggatgttt caacaagctt ggatcctcga agagaagggg taataacaca   1200
cttttttaac atttttaaca caaatttttag ttatttaaaa atttattaaa aaatttaaaa   1260
taagaagagg aactctttaa ataaatctaa cttacaaaat ttatgattht taataagtht   1320
tcaccaataa aaaatgtcat aaaaatattg taaaaagtat attatcaata ttctctttat   1380
gataaataaa aagaaaaaaaa aaataaaagt taagtgaaaa tgagattgaa gtgacttttag   1440
gtgtgtataa atatatcaac cccgccaaca atttatttaa tccaaatata ttgaagtata   1500
ttattccata gcctttatth atttatatat ttattatata aaagctttat ttgttctagg   1560
ttgttcatga aatatttttt tggttttatc tccgttgtaa gaaaatcatg tgetttgtgt   1620
cgccactcac tattgcagct ttttcatgca ttggtcagat tgacggttga ttgtattht   1680
gttttttatg gttttgtgth atgacttaag tcttcatctc tttatctctt catcaggtht   1740
gatgthtacc taatatgthc catgggtaca tgcatgthta aattagthgg ccaactthgt   1800
tgtgaacgat agaatttttt ttatattaag taaactatth ttatattatg aaataataat   1860
aaaaaaaaata ttttatcatt attaacaaaa tcatattatg taattthgth actctataat   1920
aaaagaaata ctgtaacatt cacattacat ggtaacatct tccaccctt tcattthgth   1980
tttthttgat gacttttttt cttgthttaa tttatthccc tctthttaa tthggaatac   2040
attatcatca tatataaact aaaatactaa aaacagthatt acacaaatga taaataataa   2100

```

-continued

---

cacaaatatt tataaatcta gctgcaatat atttaaacta gctatatcga tattgtaaaa	2160
taaaactagc tgcattgata ctgataaaaa aatatcatgt gctttctgga ctgatgatgc	2220
agtatacttt tgacattgcc tttattttat ttttcagaaa agctttctta gttctgggtt	2280
cttcattatt tgtttcccat ctccattgtg aattgaatca tttgcttctg gtcacaaata	2340
caatttagnt aggtacatgc attggtcaga ttcacggttt attatgcat gacttaagtt	2400
catggtagta cattacctgc cacgcacgca ttatattggg tagatttgat aggcaaat	2460
ggttgcaac aatataaata taataatgt ttttatatta cgaaataaca gtgatcaaaa	2520
caaacagttt tatctttatt aacaagattt tgtttttgtt tgatgacggt ttttaagtt	2580
tacgctttcc cccttctttt gaatttagaa cactttatca tcataaaatc aaatactaaa	2640
aaaattacat atttcataaa taataacaca aatattttta aaaaatctga aataaatg	2700
aacaatatta catattatca cgaaaattca ttaataaaaa tattatataa ataaaatgta	2760
atagtagtta tatgtaggaa aaaagtactg cacgcataat atatacaaaa agattaaaat	2820
gaactattat aaataataac actaaattaa tgggtaatca tatcaaaaata atgaaaaagt	2880
aaataaaaatt tgtaattaac ttctatatgt attacacaca caaataataa ataatagtaa	2940
aaaaaattat gataaatatt taccatctca taagatattt aaaataatga taaaaatata	3000
gattattttt tatgcaacta gctagccaaa aagagaacac gggatatatat aaaaagagta	3060
cctttaaatt ctactgtact tcctttatc ctgacgtttt tatatcaagt ggacatacgt	3120
gaagatttta attatcagtc taaatatttc attagcactt aatacttttc tgttttattc	3180
ctatcctata agtagtcccg attctcccaa cattgcttat tcacacaact aactaagaaa	3240
gtcttcata gcccccaag cggcccatgg cctcctccga ggacgcatc aaggagtca	3300
tgcgcttcaa ggtgcgatg gagggctccg tgaacggcca cgagttcgag atcgagggcg	3360
agggcgaggg cgcctctac gagggcacc agaccgcaa gctgaagggt accaagggcg	3420
gccccctgcc cttegcctgg gacatcctgt cccccagtt ccagtaeaggc tccaaggtgt	3480
acgtgaagca ccccgccgac atccccgact acaagaagct gtccttcccc gagggcttca	3540
agtgggagcg cgtgatgaac ttcgaggacg gcggcgtggt gaccgtgacc caggactcct	3600
ccctgcagga cggctccttc atctacaagg tgaagttcat cggcgtgaac ttcccctccg	3660
acggccccgt aatgcagaag aagactatgg gctgggaggg ctccaccgag cgctgtacc	3720
cccgcgacgg cgtgctgaag ggcgagatcc acaaggcctt gaagotgaag gacggcggcc	3780
actacctggt ggagttcaag tccatctaca tggccaagaa gccctgtcag ctgcccggct	3840
actactacgt ggactccaag ctggacatca cctcccacaa cgaggactac accatcgtgg	3900
agcagtacga gcgcccgag ggcgcccacc acctgttctt gtagcgccg gccgacac	3960
aagtgtgaga gtactaaata aatgctttgg ttgtacgaaa tcattacact aaataaata	4020
atcaaagctt atatatgctt tccgctaagg ccgaatgcaa agaaattggt tctttctctg	4080
tatcttttgc cacttttact agtacgtatt aattactact taatcatctt tgtttacggc	4140
tcattatata cgtcgacggc gcgcccctct agagggccca attcgcccta tagtgagtcg	4200
tattacaatt cactggccgt cgttttacia cgtcgtgact gggaaaaacc tggcgttacc	4260
caacttaatc gccttgacgc acatccccct ttcgccagct ggcgtaatag cgaagaggcc	4320
cgcaccgac gcccttccca acagttgcgc agcctatacg tacggcagtt taaggtttac	4380
acctataaaa gagagagccg ttatcgtctg tttgtggatg tacagagtga tattattgac	4440
acgcccggggc gacggatggt gatccccctg gccagtgcac gtctgctgac agataaagtc	4500

-continued

---

tcccgtgaac	tttaccgggt	ggtgcatatc	ggggatgaaa	gctggcgcac	gatgaccacc	4560
gatatggcca	gtgtgccggg	ctccgttata	ggggaagaag	tggtgatctc	cagccaccgc	4620
gaaaatgaca	tcaaaaacgc	cattaacctg	atggtctggg	gaatataaat	gtcaggcatg	4680
agattatcaa	aaaggatcct	cacctagatc	cttttcacgt	agaaagccag	tccgcagaaa	4740
cggtgctgac	cccggatgaa	tgtagctac	tgggctatct	ggacaaggga	aaacgcaagc	4800
gcaaagagaa	agcaggtagc	ttgcagtggg	cttacatggc	gatagctaga	ctgggcgggt	4860
ttatggacag	caagcgaacc	ggaattgcc	gctggggcgc	cctctggtaa	ggttgggaag	4920
ccctgcaaag	taaactggat	ggctttcttg	ccgccaagga	tctgatggcg	caggggatca	4980
agctctgac	aagagacagg	atgaggatcg	tttcgcatga	ttgaacaaga	tggattgcac	5040
gcaggttctc	cggccgcttg	ggtggagagg	ctattcggct	atgactgggc	acaacagaca	5100
atcggctgct	ctgatgcgcg	cgtgttcocg	ctgtcagcgc	aggggcgccc	ggttcttttt	5160
gtcaagaccg	acctgtccgg	tgccctgaat	gaactgcaag	acgaggcagc	gcggtatcgc	5220
tggctggcca	cgaagggcgt	tccttgccga	gctgtgctcg	acgttgctac	tgaagcggga	5280
agggactggc	tgctattggg	cgaagtgcg	gggcaggatc	tcctgtcacc	tcacctgct	5340
cctgccgaga	aagtatccat	catggctgat	gcaatgcggc	ggctgcatac	gcttgatccg	5400
gtacctgcc	cattcgacca	ccaagcga	catcgcatcg	agcgagcagc	tactcggatg	5460
gaagccggtc	ttgtcgatca	ggatgatctg	gacgaagagc	atcaggggct	cgcgccagcc	5520
gaactgttcg	ccaggctcaa	ggcgagcatg	cccacgcggc	aggatctcgt	cgtgacccat	5580
ggcgatgcct	gcttgccgaa	tatcatgggtg	gaaaatggcc	gcttttctgg	attcatcgac	5640
tgtggccggc	tgggtgtggc	ggaccgctat	caggacatag	cggtggctac	ccgtgatatt	5700
gctgaagagc	ttggcggcga	atgggctgac	cgttctctcg	tgctttacgg	tatcgccgct	5760
cccgatctgc	agcgcacgcg	cttctatcgc	cttcttgacg	agttctctcg	aattattaac	5820
gcttacaatt	tcctgatgcg	gtatttctc	cttacgcacc	tgtgoggtat	ttcacaccgc	5880
atcagggtggc	acttttcggg	gaaatgtgcg	cggaaacctc	atgtgttat	ttttctaaat	5940
acattcaaat	atgtatccgc	tcattgagaca	ataacctga	taaatgcttc	aataatagca	6000
cgtgaggagg	gccaccatgg	ccaagttgac	cagtgcgctt	ccggtgctca	ccgcgcgca	6060
cgtagccgga	gcggtcgagt	tctggaccga	cggctcggg	ttctccggg	acttcgtgga	6120
ggacgacttc	gcccgtgtgg	tcggggacga	cgtgacctg	ttcatcagcg	cggtccagga	6180
ccagggtggg	ccggacaaca	ccctggcctg	ggtgtgggtg	cgcggcctgg	acgagctgta	6240
cgcagagtgg	tcggaggtcg	tgtccacgaa	cttccgggac	gcctccgggc	cggccatgac	6300
cgagatcggc	gagcagccgt	gggggcccga	gttcgcccctg	cgggacctcg	ccggcaactg	6360
cgtgcacttc	gtggccgagg	agcaggactg	acacgtgcta	aaacttcatt	tttaatttaa	6420
aaggatctag	gtgaagatcc	tttttgataa	tctcatgacc	aaaatccctt	aacgtgagtt	6480
ttcgttccac	tgagcgtcag	accccgtaga	aaagatcaaa	ggatcttctt	gagatccttt	6540
ttttctgcgc	gtaactctgct	gcttgcaaac	aaaaaaacca	ccgctaccag	cggtgggttg	6600
tttgccggat	caagagctac	caactctttt	tccgaaggta	actggcttca	gcagagcgca	6660
gataccaaat	actgttcttc	tagttagacc	gtagttaggc	caccacttca	agaactctgt	6720
agcaccgcct	acatacctcg	ctctgcta	cctgttacca	gtggctgctg	ccagtgccga	6780
taagtcgtgt	cttaccgggt	tggactcaag	acgatagtta	ccggataagg	cgcagcggtc	6840

-continued

---

gggctgaacg ggggggttcgt gcacacagcc cagcttggag cgaacgacct acaccgaact	6900
gagataccta cagcgtgagc tatgagaaaag cgccacgctt cccgaaggga gaaaggcgga	6960
caggtatccg gtaagcggca gggtcggaac aggagagcgc acgagggagc ttccaggggg	7020
aaacgcctgg tatctttata gtcctgtcgg gtttcgccac ctctgacttg agcgtcgatt	7080
tttgtgatgc tcgtcagggg ggcgggagcct atggaaaaac gccagcaacg cggccttttt	7140
acggttcctg gccttttgcg ggccttttgc tcacatgttc tttcctgcgt tatccctga	7200
ttctgtgat aaccgtatta ccgccttga gtgagctgat accgctcgc gcagccgaac	7260
gaccgagcgc agcaggtcag tgagcgagga agcgggaagag cgccaatac gcaaaccgcc	7320
tctccccgcg cgttggccga ttcattaatg cagctggcac gacaggttcc cgcactggaa	7380
agcgggcagt gagcgcaacg caattaatgt gagttagctc actcattagg caccocaggc	7440
tttacacttt atgcttcocg ctctgatgtt gtgtggaatt gtgagcggat aacaatttca	7500
cacaggaaac agctatgacc atgattacgc caagctattt aggtgacgcg ttagaatact	7560
caagctatgc atcaagcttg gtaccgagct cggatccact agtaacggcc gccagtgtgc	7620
tggaattcag gggcgcgccc caatttatta tcatctattt cctgacattt taatccatcc	7680
acctatgtca aaaacttata gaaaaatgtca acttccaaac aaaacataat tgaacttcgc	7740
aaataaattc ttaataatat taaaaaatgt tacttaatta tttcttcaac cccattttcc	7800
gcgcgtagcg cggacaaaaga ctctagttaa atatagaagt ttccgattct catcgtataa	7860
aacggtgact ttggcgggct ttcattgtgt acaaatgggt ttaacaaacc actgcctagt	7920
cgtttagtgt agaatcagcg catggaactc cgattggagc gtgactttca cgtgccggag	7980
gcccaccacc acagcggggc ttacgctcta agaactcgc ccacggtttt ctctcatctc	8040
cccccgccaa gtgtctcctc cgttcgocac ttctcatcat gttacagggga ccataaaaat	8100
ggcgtatttc ttcagccccg ggtataaata cacacatgat cctgtggtgg gttcttcac	8160
aagttacatc tccttctggg ttttgtattg caagtgtttg tattttttgc ctccgagaga	8220
aaatcgcggc cgcattgctg ctgctcccag tgtgaggacg tttactcggg ccgaggtttt	8280
gaatgccgag gctctgaatg agggcaagaa ggatgccgag gcaccctctc tgatgatcat	8340
cgacaacaag gtgtacgatg tccgcgagtt cgtccctgat catcccggtg gaagtgtgat	8400
tctcacgcac gttggcaagg acggcactga cgtctttgac acttttcacc ccgaggtcgc	8460
ttgggagact cttgccaact tttacgttgg tgatattgac gagagcgacc gcgatatcaa	8520
gaatgatgac tttcggcgcg aggtccgcaa gctcgtgacc ttgttccagt ctcttggtta	8580
ctacgattct tccaaggcat actacgctt caaggtctcg ttcaacctct gcactcgggg	8640
tttgtcagc gtcattgtgg ccaagtgggg ccagacctcg accctcgcca acgtgctctc	8700
ggctgcgctt ttgggtctgt tctggcagca gtgcccggatgg ttggctcagc actttttgca	8760
tcaccaggtc ttccaggacc gtttctgggg tgatcttttc ggcgcttct tgggaggtgt	8820
ctgccagggc ttctcgtcct cgtggtgga ggacaagcac aacctcacc acgcccgcc	8880
caacgtccac ggcgaggatc ccgacattga caccaccct ctgttgacct ggagtgagca	8940
tgcgttgag atgttctcgg atgtcccaga tgaggagctg acccgcatgt ggtcgcgttt	9000
catggtcctg aaccagacct ggttttactt cccattctc tcgtttgccc gtctctcctg	9060
gtgcctccag tccattctct ttgtgtgcc taacggtcag gcccaacaagc cctcgggcgc	9120
gcgtgtgccc atctcgttgg tcgagcagct gtcgcttgcg atgcactgga cctggtacct	9180
cgccaccatg ttctgttca tcaaggatcc cgtcaacatg ctggtgtact ttttgtgtc	9240

-continued

---

```

gcaggcggtg tgcggaaact tgttggcgat cgtgttctcg ctcaaccaca acggtatgcc 9300
tgtgatctcg aaggaggagg cggtcgatat ggatttcttc acgaagcaga tcatcacggg 9360
tcgtgatgtc caccggggtc tatttgccaa ctggttcacg ggtggattga actatcagat 9420
cgagcaccac ttgttcctt cgatgcctcg ccacaacttt tcaaagatcc agcctgctgt 9480
cgagaccctg tgcaaaaagt acaatgtccg ataccacacc accggtatga tcgagggaac 9540
tcgagaggtc tttagccgtc tgaacgaggt ctccaagget gcctccaaga tgggtaaggc 9600
gcagtaagc 9609

```

```

<210> SEQ ID NO 61
<211> LENGTH: 19404
<212> TYPE: DNA
<213> ORGANISM: Artificial sequence
<220> FEATURE:
<223> OTHER INFORMATION: Vector ARAL080
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: (17594)..(17594)
<223> OTHER INFORMATION: n is a, c, g, or t

```

```

<400> SEQUENCE: 61

```

```

cgcgctcga gtggcggat cccccgggct gcaggaattc actggccgtc gttttacaac 60
gtcgtgactg ggaaaacctt ggcggtaccc aacttaatcg ccttgacga catccccctt 120
tcgccagctg gcgtaatagc gaagaggccc gcaccgatcg cccttccaa cagttgcgca 180
gcctgaatgg cgaatggatc gatccatcgc gatgtacctt ttgttagtca gcctctcgat 240
tgctcatcgt cattacacag taccgaagtt tgatcgatct agtaacatag atgacaccgc 300
gcgcgataat ttatcctagt ttgcgcgcta tttttgttt tctatcgcgt attaaatgta 360
taattgctgg actctaata taaaaacca tctcataaat aacgtcatgc attacatggt 420
aattattaca tgcttaacgt aattcaacag aaattatatg ataatcatcg caagaccggc 480
aacaggattc aatcttaaga aactttattg ccaaatgttt gaacgatctg cttcgacgca 540
ctccttcttt actccaccat ctgcctctta ttgaaaacgt gggtagcacc aaaacgaatc 600
aagtcgctgg aactgaagtt accaatcacg ctggatgatt tgccagttgg attaactctg 660
cctttccccg catgaataat attgatgaat gcatgcgtga ggggtagtcc gatggtggca 720
atagctgcaa ttgccgcgac atcctccaac gagcataatt cttcagaaaa atagcgatgt 780
tccatgttgt cagggcatgc atgatgcacg ttatgaggtg acggtgctag gcagatttcc 840
ctcaaagttt catagtcagt atcatattca tcattgcatt cctgcaagag agaattgaga 900
cgcaatccac acgctgcggc aaccttccgg cgttcgtggt ctatttgctc ttggacgttg 960
caaacgtaag tgttggatcg atccggggtg ggccaagaac tccagcatga gatccccgcg 1020
ctggaggatc atccagccgg cgtcccggaa aacgattccg aagcccaacc tttcatagaa 1080
ggcggcggtg gaatcgaat ctcgtgatgg caggttgggc gtcgcttggc cggtcatttc 1140
gaaccccaga gtcccgtca gaagaactcg tcaagaaggc gatagaaggc gatgcgctgc 1200
gaatcgggag cggcgatacc gtaaagcacg aggaagcggg cagcccattc gccgccaagc 1260
tcttcagcaa tatcacgggt agccaacgct atgtcctgat agcgggtccc cacaccagc 1320
cggccacagt cgatgaatcc agaaaagcgg ccattttcca ccatgatatt cggcaagcag 1380
gcatcgccat gggtcacgac gagatcctcg ccgtcgggca tgcgcgctt gagcctggcg 1440
aacagttcgg ctggcgcgag ccctgatgc tcttcgtcca gatcctctg atcgacaaga 1500

```

-continued

---

cgggcttcca	tccgagtacg	tgctcgctcg	atgcgatggt	tcgcttgggtg	gtcgaatggg	1560
caggtagccg	gatcaagcgt	atgcagccgc	cgcatctgat	cagccatgat	ggatactttc	1620
tcggcaggag	caaggtgaga	tgacaggaga	tcttgccccg	gcacttcgcc	caatagcagc	1680
cagtcccttc	ccgcttcagt	gacaacgtcg	agcacagctg	cgcaaggaac	gcccgctcgtg	1740
gccagccacg	atagccgcgc	tgccctcgtcc	tcgagttcat	tcagggcacc	ggacaggctcg	1800
gtcttgacaa	aaagaaccgg	gcgcccctgc	gctgacagcc	ggaacacggc	ggcatcagag	1860
cagccgattg	tctgtttgtc	ccagtcatag	ccgaatagcc	tctccacca	agcggccgga	1920
gaacctgcgt	gcaatccatc	ttgttcaatc	atgcgaaacg	atccccgcaa	gcttggagac	1980
tggtgatttc	agcgtgtcct	ctccaaatga	aatgaacttc	cttatataga	ggaagggctc	2040
tgcaaggat	agtgggattg	tcgctcatcc	cttacgtcag	tggagatata	acatcaatcc	2100
acttgctttg	aagacgtgg	tggaaactct	tctttttcca	cgatgctcct	cgtgggtggg	2160
ggctccatct	tgggaccact	gtcggcagag	gcattctcaa	cgatggcctt	tcctttatcg	2220
caatgatggc	attttagtag	gccaccttcc	ttttccacta	tcttcacaat	aaagtgcagc	2280
atagctgggc	aatggaatcc	gaggagggtt	ccggatatta	ccctttgttg	aaaagtctca	2340
attgcccttt	ggctctctga	gactgtatct	ttgatatttt	tggagtagac	aagcgtgctg	2400
tgctccacca	tggtgacgaa	gattttcttc	ttgtcattga	gtcgtaaagag	actctgtatg	2460
aactgtctgc	cagctcttac	ggcgagttct	gttaggtcct	ctatttgaat	ctttgactcc	2520
atggcctttg	attcagtggt	aactaccttt	ttagagactc	caatctctat	tacttgctct	2580
ggtttgtgaa	gcaagccttg	aatcgtccat	actggaatag	tacttctgat	cttgagaaat	2640
atatctttct	ctgtgttctt	gatgcagtta	gtcctgaatc	ttttgactgc	atctttaacc	2700
ttcttgggaa	ggatattgat	ctcctggaga	ttattgctcg	ggtagatcgt	cttgatgaga	2760
cctgctgcgt	aagcctctct	aaccatctgt	gggttagcat	tctttctgaa	attgaaaagg	2820
ctaactctct	cattatcagt	ggatgaacatg	gtatcgtcac	cttctccgctc	gaacttcctg	2880
actagatcgt	agagatagag	gaagtcgtcc	attgtgatct	ctggggcaaa	ggagatctga	2940
attaattcga	tatggtggat	ttatcacaaa	tgggacccgc	cgccgacaga	ggtgtgatgt	3000
taggccagga	ctttgaaaat	ttgcgcaact	atcgtatagt	ggccgacaaa	ttgacccgga	3060
gttgacagac	tgccatagcat	ttgagtgaat	tatgtgaggt	aatgggctac	actgaattgg	3120
tagctcaaac	tgctcagtatt	tatgtatatg	agtgtatatt	ttcgcataat	ctcagaccaa	3180
tctgaagatg	aaatgggtat	ctgggaatgg	cgaaatcaag	gcatacgatcg	tgaagtttct	3240
catctaagcc	cccatttggg	cgtgaatgta	gacacgtcga	aataaagatt	tccgaattag	3300
aataatttgt	ttattgcttt	cgccataaaa	tacgacggat	cgtaatttgt	cgttttatca	3360
aaatgtactt	tcattttata	ataacgctgc	ggacatctac	atttttgaat	tgaaaaaaaa	3420
ttggtaatta	ctctttcttt	ttctccatat	tgaccatcat	actcattgct	gatccatgta	3480
gatttcccg	acatgaagcc	atttacaatt	gaatataatc	tgccgcccgt	gccgctttgc	3540
acccgggtga	gcttgcattg	tggtttctac	gcagaactga	gccgggttagg	cagataattt	3600
ccattgagaa	ctgagccatg	tgacaccttc	ccccaacacg	gtgagcgacg	gggcaacgga	3660
gtgatccaca	tgggactttt	aaacatcatc	cgctcgatgg	cgttgcgaga	gaagcagctc	3720
atccgtgaga	tcagccgacg	caccgggcag	gcgcgcaaca	cgatcgcaaa	gtatttgaac	3780
gcaggtacaa	tcgagccgac	gttcacgcgg	aacgaccaag	caagctagct	ttaatgcggt	3840
agtttatcac	agttaaattg	ctaacgcagt	caggcacctg	gtatgaaatc	taacaatgct	3900

-continued

---

ctcatcgtea	tcctcggcac	cgccacctg	gatgctgtag	gcataggctt	ggttatgccc	3960
gtactgccc	gcctcttgcg	ggatctctc	cattccgaca	gcatcgccag	tcaactatggc	4020
gtgctgctag	cgctatatgc	gttgatgcaa	tttctatgcg	caccogttct	cgagagcactg	4080
tccgaccgct	ttggccgccc	cccagtcctg	ctcgtctcgc	tacttgagagc	caactatcgac	4140
tacgcgatca	tggcgaccac	accgctcctg	tggtccaacc	cctccgctgc	tatagtgcag	4200
tccgcttctg	acgttcagtg	cagccgtctt	ctgaaaacga	catgtcgcac	aagtcctaag	4260
ttacgcgaca	ggctgcgccc	ctgccccttt	cctggcgctt	tcttgctcgcg	tgttttagtc	4320
gcataaagta	gaatacttgc	gactagaacc	ggagacatta	cgccatgaac	aagagcgcgg	4380
ccgctggcct	gctgggtat	gcccgcgca	gcaccgacga	ccaggacttg	accaaccaac	4440
gggcccgaact	gcacgcggcc	ggctgcacca	agctgttttc	cgagaagatc	accggcacca	4500
ggcgcgacgg	cccggagctg	gccaggatgc	ttgaccacct	acgcccctggc	gacgttgga	4560
cagtgaccag	gctagaccgc	ctggcccga	gcaccgcga	cctactggac	attgcccagc	4620
gcatccagga	ggcccggcgg	ggcctgcgta	gctggcaga	gcccgtggcc	gacaccacca	4680
cgcccggccc	ccgcatggtg	ttgaccgtgt	tcgcccggcat	tgccgagttc	gagcgttccc	4740
taatcatcga	ccgcaccogg	agcgggccc	aggcccga	ggcccagggc	gtgaagtgtg	4800
gccccgccc	tacctcacc	ccggcacaga	tcgcccgcgc	ccgagagctg	atcgaccagg	4860
aaggcccgcac	cgtgaaagag	gcccgtgcac	tgcttggcgt	gcatcgctcg	accctgtacc	4920
gcgcacttga	gcccagcagc	gaagtgcgc	ccaccgagcc	caggcggcgc	ggtgccttcc	4980
gtgaggagc	attgaccgag	gcccagccc	tgcccggccc	cgagaatgaa	cgccaagagg	5040
aacaagcatg	aaaccgcacc	aggacggcca	ggacgaaccg	ttttcatta	ccgaagagat	5100
cgaggcggag	atgatcggg	ccgggtacgt	gttcgagccc	cccgcgcacg	tctcaaccgt	5160
gcccgtgcat	gaaatcctgg	ccggtttgtc	tgatgccaag	ctggcggcct	ggcccggcag	5220
cttgcccgt	gaagaaaccg	agcggcccgg	tctaaaagg	tgatgtgtat	ttgagtaaaa	5280
cagcttgcgt	catcggctc	ctgcgtatat	gatgcgatga	gtaaataaac	aaatacgcaa	5340
gggaacgcat	gaagtatc	ctgtacttaa	ccagaaaggc	gggtcaggca	agacgaccat	5400
cgcaaccat	ctagcccgg	ccctgcaact	cgcccggccc	gatgttctgt	tagtcgattc	5460
cgatcccag	ggcagtgccc	gcgattgggc	ggcctgccc	gaagatcaac	cgctaaccgt	5520
tgccggcatc	gaccgcccga	cgattgaccg	cgacgtgaag	gccatcggcc	ggcgcgactt	5580
cgtagtgatc	gacggagcgc	cccaggcggc	ggacttggct	gtgtcccga	tcaaggcagc	5640
cgacttctgt	ctgattccgg	tcagcccag	cccttacgac	atatgggcca	ccgcccacct	5700
ggggagctg	gttaagcagc	gcattgaggt	cacggatgga	aggctacaag	cgccctttgt	5760
cgtgtcggg	gcatcaaaag	gcacgcgcac	cgccgggtgag	gttgcccagg	cgctggccc	5820
gtacgagctg	cccattcttg	agtcctcgat	cacgcagcgc	gtgagctacc	caggcactgc	5880
cgcccgggc	acaaccgttc	ttgaatcaga	accgagggc	gacgctgccc	gagaggtcca	5940
ggcgtggcc	gctgaaatta	aatcaaaact	catttgagtt	aatgaggtaa	agagaaaatg	6000
agcaaaagca	caaacacgct	aagtgcggc	cgccggagcg	cacgcagcag	caaggctgca	6060
acgttggcca	gcctggcaga	caccccagcc	atgaagcggg	tcaactttca	ggtgcccggc	6120
gaggatcaca	ccaagctgaa	gatgtacgg	gtacgccaag	gcaagaccat	taccgagctg	6180
ctatctgaat	acatcgcgca	gctaccagag	taaatgagca	aatgaataaa	tgagtagatg	6240

-continued

---

aattttagcg gctaaaggag gcgcatgga aatcaagaa caaccaggca cggacgccgt	6300
ggaatgcccc atgtgtggag gaaccggcgg ttggccaggc gtaagcggct gggttgtctg	6360
cggccctgc aatggcactg gaaccccaa gcccaggaa tcggcgtgag cggtcgcaa	6420
ccatccggcc cggtaacaaat cggcggcggc ctgggtgatg acctggtgga gaagttgaag	6480
gccgcgcagg cggcccagcg gcaacgcatac gaggcagaag cacgcccgg tgaatcgtgg	6540
caagcggcgg ctgatcgaat ccgcaagaa tcccggaac cgccggcagc cggtgccggc	6600
tcgattagga agccgccccaa gggcgacgag caaccagatt ttttcgttcc gatgctctat	6660
gacgtgggca ccccgcatag tcgcagcatac atggacgtgg ccgttttccg tctgtcgaag	6720
cgtgaccgac gagctggcga ggtgatccgc tacgagcttc cagacgggca cgtagaggtt	6780
tccgcagggc cggccggcat ggccagtgtg tgggattacg acctggtact gatggcgggt	6840
tcccataaa ccgaatccat gaaccgatac cgggaaggga agggagacaa gcccgccgc	6900
gtgttccgtc cacacgttgc ggacgtactc aagttctgcc ggcgagccga tggcgaaaag	6960
cagaaagacg acctggtaga aacctgcatt cggttaaaca ccacgcacgt tgccatgcag	7020
cgtacgaaga aggccaagaa cggccgcctg gtgacgggat ccgaggggta agccttgatt	7080
agccgctaca agatcgtaaa gagcgaacc gggcggcggc agtacatcga gatcgagcta	7140
gctgattgga tgtaccgca gatcacagaa ggcaagaacc cggacgtgct gacggttcac	7200
cccgattact ttttgatcga tcccgcatc gcccgtttcc tctaccgctt gccacgccgc	7260
gccgcaggca aggcagaagc cagatgggtt ttcaagacga tctacgaacg cagtggcagc	7320
gccggagagt tcaagaagtt ctgtttcacc gtgcgcaagc tgatcgggtc aatgacctg	7380
cgggagtacg atttgaagga ggaggcgggg caggctggcc cgatcctagt catgcctac	7440
cgcaacctga tcgagggcga agcatccgcc ggttcctaata gtacggagca gatgctaggg	7500
caaattgccc tagcagggga aaaaggtcga aaaggtctct ttctgtgga tagcacgtac	7560
attgggaacc caaagccgta cattgggaac cggaaacctg acattgggaa cccaaagccg	7620
tacattggga accggtcaca catgtaagt actgatataa aagagaaaa aggcgatttt	7680
tccgcctaaa actctttaa acttattaaa actcttaaaa cccgcctggc ctgtgcataa	7740
ctgtctggcc agcgcacagc cgaagagctg caaaaagcgc ctacccttcg gtcgctgcgc	7800
tccctacgcc ccgccccttc gcgtcggcct atcggcggccg ctggccgctc aaaaatggct	7860
ggcctacggc caggcaatct accagggcgc ggacaagccg cgcgctcgc actcgaccgc	7920
cggcgcccaac atcaaggcac cctgcctcgc gcgtttcggg gatgacgggtg aaaaacctctg	7980
acacatgcag ctcccggaga cggtcacagc ttgtctgtaa gcggatgccg ggagcagaca	8040
agcccgtcag ggcgcgtcag cgggtgttgg cgggtgtcgg ggcgcagcca tgaccagtc	8100
acgtagcgat agcggagtgat atactggctt aactatgcgg catcagagca gattgtactg	8160
agagtgcacc atatcgggtg tgaataaccg cacagatgcy taaggagaaa atacgcatac	8220
aggcgtctt cgccttcctc gctcactgac tcgctgcgct cggtcgttcg gctgcggcga	8280
gcggtatcag ctcactcaaa ggcggttaata cggttatcca cagaatcagg ggataacgca	8340
ggaaagaaca tgtgagcaaa aggccagca aagccagga accgtaaaaa ggccgcgttg	8400
ctggcgtttt tccataggct ccgccccct gacgagcatac acaaaaatcg acgctcaagt	8460
cagaggtggc gaaacctgac aggactataa agataaccagg cgtttcccc tggaagctcc	8520
ctcgtgcgct ctctgttcc gacctgccc cttaccggat acctgtccc ctttctccct	8580
tcgggaagcg tggcgcttcc tcatagctca cgctgtaggt atctcagttc ggtgtaggtc	8640

-continued

---

gttcgctcca agctgggctg tgtgcaogaa cccccggtc ageccgaccg ctgogcotta 8700  
 tccgtaact atcgtcttga gtccaacccg gtaagacacg acttatcgcc actggcagca 8760  
 gccactggta acaggattag cagagcgagg tatgtaggcg gtgctacaga gttcttgaag 8820  
 tgggtggccta actacggcta cactagaagg acagtatttg gtatctgccc tctgctgaag 8880  
 ccagttacct tcggaaaaag agttggtagc tcttgatccg gcaaaaaaac caccgctggt 8940  
 agcggtaggt tttttgttg caagcagcag attacgcgca gaaaaaagg atctcaagaa 9000  
 gatccttga tctttctac ggggtctgac gctcagtggc acgaaaactc acgtaaggg 9060  
 attttggtca tgagattatc aaaaaggatc ttcacctaga tctttttaa ttaaaaatga 9120  
 agttttaaact caatctaaag tatatatgag taaacttggc ctgacagtta ccaatgctta 9180  
 atcagtgagg cacctatctc agcgatctgt ctatttcggt catccatagt tgctgactc 9240  
 cccgtcgtg agataactac gatacgggag ggcttaccat ctggcccag tgctgcaatg 9300  
 ataccgcgag acccacgctc accggctcca gatttatcag caataaacca gccagccgga 9360  
 agggccgagc gcagaagtgg tctgcaact ttatccgctc ccatccagtc tattaattgt 9420  
 tgccgggaag ctagagtaag tagttcgcca gttaatagtt tgcgcaacgt tgttgccatt 9480  
 gctacaggca tctggtgtc acgctcgtcg tttggtatgg cttcattcag ctccggttc 9540  
 caacgatcaa ggcgagttac atgatcccc atgttgtgca aaaaagcggc tagctcctc 9600  
 ggtcctccga tctgtgtcag aagtaagttg gccgcagtgt tatcactcat ggttatggca 9660  
 gcactgcata attctcttac tgcctatgcca tccgtaagat gctttctgt gactggtgag 9720  
 tactcaacca agtcattctg agaatagtg atgcggcgac cgagttgctc ttgccggcg 9780  
 tcaacacggg ataataccgc gccacatagc agaactttaa aagtgtcat cattggaaaa 9840  
 gacctgcagg gggggggggg cgtgagggtc tgcctcgtga agaaggtgtt gctgactcat 9900  
 accaggcctg aatcgcccc tcactccagcc agaaagtgag ggagccacgg ttgatgagag 9960  
 ctttgttga ggtggaccag ttggtgattt tgaacttttg ctttgcacg gaacggtctg 10020  
 cgttgtcggg aagatgcgtg atctgaccc tcaactcagc aaaagttcga tttattcaac 10080  
 aaagccgccc tcccgtcaag tcagcgtaat gctctgccag tgttacaacc aattaacca 10140  
 ttctgattag aaaaactcat cgagcatcaa atgaaactgc aatttattca tatcaggatt 10200  
 atcaatacca ttttttga aaagccggtt ctgtaatgaa ggagaaaact caccgaggca 10260  
 gttccatagg atggcaagat cctggtatcg gtctcgatt cggactcgtc caacatcaat 10320  
 acaacctatt aatttccctc cgtcaaaaat aaggttatca agtgagaaat caccatgagt 10380  
 gacgactgaa tccggtgaga atggcaaaag cttatgcatt tctttccaga cttgttcaac 10440  
 aggccagcca ttacgctcgt catcaaaatc actcgcata accaaaaccg tattcattcg 10500  
 tgattgcgcc tgagcgagac gaaataccgc atcgtgtta aaaggacaat tacaacagg 10560  
 aatcgaatgc aaccggcgca ggaacactgc cagcgcata acaatatttt cacctgaatc 10620  
 aggatattct tctaatacct ggaatgctgt tttccgggg atcgcagtgg tgagtaacca 10680  
 tgcatcatca ggagtacgga taaaatgctt gatggtcgga agaggcataa attccgctcag 10740  
 ccagtttagt ctgaccatct catctgtaac atcattggca acgctacctt tgccatggtt 10800  
 cagaaacaac tctggcgcgt cgggcttccc atacaatcga tagattgtcg cacctgattg 10860  
 cccgacatta tcgcgagccc atttatacc atataaatca gcatccatgt tggaatttaa 10920  
 tcgggcccgc gagcaagacg tttccggtg aatatggctc ataaccccc ttgtattact 10980

-continued

---

gtttatgtaa	gcagacagtt	ttattgttca	tgatgatata	ttttatctt	gtgcaatgta	11040
acacagaga	ttttgagaca	caacgtggct	ttccccccc	cccctgcagg	tcaattcggc	11100
cgatatggct	attacgaaga	aggctcgtgc	gcgaggtccc	gtgaactttc	ccacgcaaca	11160
agtgaaccgc	accgggtttg	ccggaggcca	tttcgttaaa	atgctgcagcc	atggctgctt	11220
cgtccagcat	ggcgtaatac	tgatcctcgt	cttcggctgg	cggtatatgg	ccgatgggct	11280
tcaaaagccg	ccgtggttga	accagtcctat	ccattccaag	gtagcgaact	cgaccgcttc	11340
gaagctcttc	catggctccac	gccgatgaat	gacctcggcc	ttgtaaagac	cgttgatcgc	11400
ttctgcgagg	gcgctgtcgt	gctgtcgcgc	acgcttccca	tagatggctc	gatacctgct	11460
tctgccaaac	gctcgggaata	gcgaaaggac	acgtattgaa	caccgcgac	cgagtgatgc	11520
actaggccgc	catgagcggg	acgccgatca	tgatgagcct	cctcgagggc	atcgaggaca	11580
aagcctgcat	gtgctgtccg	gctcgcgcc	catccgacaa	tgcgacgggc	gaagacgtcg	11640
atcacgaagg	ccaagctagac	gaagccctcc	caagtggcga	cataagtacg	gacatgcgca	11700
aaggctttcc	cggtttgtcg	ctgatggctc	aagagacgct	gaagcgcgat	ccgatgcgca	11760
ggcatctgtt	cgtcttcgcg	ggctcgtggc	gtggcctgat	caaggctcact	cgccgaagag	11820
ctgcatgatt	ggctcgaaac	cgagcggggg	aaattgtcgc	gcagttctcc	cgtcgcggag	11880
gcgataaatt	acatgctcaa	gcatgggat	ggcattacgt	cattcctcga	tgacggccc	11940
atctgcctga	cgaacaatgc	tgccgaacga	acgctcagag	gctatgtact	cggcaggaag	12000
tcacggctgt	ttgccggatc	ggatcgttgt	gctgaacgtg	cggcgttcat	ggcgacactg	12060
atcatgagcg	ccaagctcaa	taacatcgat	ccgcaggcct	ggcttgccga	cgtccgcgcc	12120
gaccttgccg	acgctccgat	cagcaggctt	gagcaacagc	tgccgtggaa	ctggacatcc	12180
aagacactga	gtgctcaggc	ggcctgacct	gcgcccttca	ccggataact	accccattat	12240
cgcagattgc	gatgaagcat	cagcgtcatt	cagcaatctt	gccaaagtat	gcaggctcgc	12300
gagaatcgac	gtgcgaaacc	ggctgggttc	gccaaagatc	cgcttgccga	gctggcgaac	12360
atcatgctg	ggacttcaag	aggctcagta	gaggaagaac	cggaaagggt	gcaccggaaa	12420
atatgcgttc	ctttggagag	cgctcatgg	acgtgaacaa	atcgcgccga	ccaaggatgc	12480
cacggataca	aaagctcgcg	aagctcggtc	ccgtgggtgt	tctgtcgtct	cgttgtacaa	12540
cgaatccat	tcccattccg	cgtcaagat	ggcttcccct	cggcagttca	tcagggttaa	12600
atcaatctag	ccgacttgtc	cgtgaaatg	ggctgcactc	caacagaaac	aatcaaacaa	12660
acatacacag	cgacttattc	acacgagctc	aaattacaac	ggatatatc	ctgccagtca	12720
gcatcatcac	accaaaagtt	aggcccgaat	agtttgaat	tagaaagctc	gcaattgagg	12780
tctacaggcc	aaattcgtc	ttagccgtac	aatattactc	accggtgcga	tgcccccat	12840
cgtagggtgaa	gggtgaaatt	aatgatccat	cttgagacca	caggcccaca	acagctacca	12900
gtttcctcaa	gggtccacca	aaaacgtaag	cgcttacgta	catggctgat	aagaaaaggc	12960
aatttgtaga	tgtaacatc	caacgtcgtc	ttcagggatc	gatccaatac	gcaaaccgcc	13020
tctccccgcg	cgttgccga	ttcattaatg	cagctggcgc	gacaggtttc	ccgactggaa	13080
agcgggcagc	gagcgaacg	caattaatgt	gagttagctc	actcattagg	caccccaggc	13140
tttacacttt	atgcttcgg	ctcgtatgtt	gtgtggaatt	gtgagcggat	aacaatttca	13200
cacaggaaac	agctatgacc	atgattacgc	caagettgca	tgctgcagg	tcgactctag	13260
aggatctggc	gcccccaat	ttattatcat	ctatttctg	acattttaat	ccatccacct	13320
atgtcaaaaa	cttatagaaa	atgtcaactt	ccaaacaaaa	cataattgaa	cttcgcaaat	13380

-continued

---

aaattcttaa taatattaaa aaatgttact taattatttc tcaacccca tttccgcgc 13440  
gtagcgcgga caaagactct agttaaataat agaagtttcc gattctcacc gtataaaacg 13500  
gtgactttgg cgggctttca tgtgtaacaa attggtttaa caaacactg cctagtcggt 13560  
tagtgtagaa tcagcgcacg gaactccgat tggagcgtga ctttcacgtg cgggaggccc 13620  
accaccacag cgggctttac gctctaagaa tctcgcacc gggtttcttc atctccccc 13680  
cgccaagtgt ctccctcggt cgcaacttct catcatgtta cagggaccat aaaaatggcg 13740  
tatttcttca gccccgggta taaatacaca catgatcctg tgggtgggtc ttcacacaagt 13800  
tacatctcct tctggttttt gtattgcaag tgtttgtatt tttgcctcc gagagaaaat 13860  
cgcgccgca tggctgctgc tcccagtggt aggacgttta ctccggccga gggtttgaaat 13920  
gcccaggctc tgaatgaggg caagaaggat gccgagcac cctcttgat gatcatcgac 13980  
aacaagtggt acgatgtccg cgagttctc cctgatcacc ccggtggaag tgtgattctc 14040  
acgcacgttg gcaaggacgg cactgacgtc tttgacactt ttcacccga ggctgcttgg 14100  
gagactcttg ccaactttta cgttggtgat attgacgaga gcgaccgca tatcaagaat 14160  
gatgactttg cggccgaggt ccgcaagctg cgtacctgt tccagtctct tggttactac 14220  
gattcttcca aggcatacta cgcttcaag gtctcgttca acctctgcat ctggggtttg 14280  
tcgacggtea tttgggcaa gtggggccag acctcgacc tcgccaacgt gctctcggt 14340  
gcgcttttgg gtctgtctg gcagcagtc ggatgggtgg ctccagactt tttgatcac 14400  
caggtcttcc aggaccgtt ctgggtgat ctttccgcg cctcttggg aggtgtctgc 14460  
cagggtctct cgtctcgtg gtggaaggac aagcacaaca ctccaccgc ccccccaac 14520  
gtccacggcg aggatccga cattgacacc caccctctgt tgacctggag tgagcatgcg 14580  
ttggagatgt tctcgatgt cccagatgag gagctgacc gcattgtggt gcgtttcatg 14640  
gtcctgaacc agacctggt ttaactcccc attctctcgt ttgcccgtct ctctgggtgc 14700  
ctccagtcca ttctcttctg gctgcctaac ggtcagggcc acaagccctc gggcgcgcgt 14760  
gtgccatct cgttggctga gcagctgtc cttgcgatgc actggacctg gtacctcgc 14820  
accatgttcc tgttcatcaa ggatccgtc aacatgctgg tgtactttt ggtgtcgcag 14880  
gcggtgtgcg gaaactgtt ggcgatcgtt ttctcgtca accacaacgg tatgcctgtg 14940  
atctcgaagg aggaggcgt cgatatggat ttcttcaaga agcagatcat caccggctct 15000  
gatgtccacc cgggtctatt tgccaactgg ttcacgggtg gattgaaacta tcagatcgag 15060  
caccactgt tcccttcgat gcctcgcacc aacttttcaa agatccagcc tgctgtcgag 15120  
acctgtgca aaaagtacaa tgtccgatac cacaccaccg gtatgatcga gggaaactgca 15180  
gaggtcttta gccgtctgaa cgaggtctcc aaggtgcct ccaagatggg taaggcgcag 15240  
taagcggcgg caagtatgaa ctaaaatgca tgtagggtga agagctcatg gagagcatgg 15300  
aatattgtat ccgaccatgt aacagtataa taactgagct ccatctcact tcttctatga 15360  
ataaacaagg gatgtatga tatattaaca ctctatctat gcaccttatt gttctatgat 15420  
aaatttctc ttattattat aaatcatctg aatcgtgacg gcttatggaa tgcttcaaat 15480  
agtacaaaa caaatgtgta ctataagact ttctaacaa ttctaacctt agcattgtga 15540  
acgagacata agtgtaaga agacataaca attataatgg aagaagtttg tctccattta 15600  
tatattatat attaccact tatgtattat attaggatgt taaggagaca taacaattat 15660  
aaagagagaa gtttgatcc atttatatat tatatactac ccatttatat attatactta 15720

-continued

---

```

tccacttatt taatgtcttt ataaggtttg atccatgata tttctaatat tttagttgat 15780
atgtatatga aagggtacta tttgaactct cttactctgt ataaaggttg gatcatcctt 15840
aaagtgggtc tattaattt tattgcttct tacagataaa aaaaaatta tgagttggtt 15900
tgataaaata ttgaaggatt taaaataata ataaataaca tataatatat gtatataaat 15960
ttattataat ataacattta tctataaaaa agtaaatatt gtcataaatc tatacaatcg 16020
tttagccttg ctggacgaat ctcaattatt taaacgagag taaacatatt tgactttttg 16080
gttatttaac aaattattat ttaacactat atgaaatfff tttttttatc agcaaagaat 16140
aaaattaaat taagaaggac aatggtgtcc caatccttat acaaccaact tccacaagaa 16200
agtcaagtea gagacaacaa aaaaacaagc aaaggaaatt ttttaatttg agttgtcttg 16260
tttgctgcat aatttatgca gtaaaacact acacataacc ctttagcag tagagcaatg 16320
gttgaccgtg tgcttagctt cttttatfff atttttttat cagcaaagaa taaataaaat 16380
aaaatgagac acttcagga tgtttcaaca agcttgatc ctggaagaga agggttaata 16440
acacactfff ttaacatttt taacacaaat tttagttatt taaaaattta ttaaaaaatt 16500
taaaataaga agaggaactc tttaaataaa tctaacttac aaaatttatg atttttaata 16560
agttttcacc aataaaaaat gtcataaaaa tatgttaaaa agtatattat caatattctc 16620
tttatgataa ataaaaagaa aaaaaaaata aaagttaagt gaaaatgaga ttgaagtgac 16680
tttaggtgtg tataaatata tcaacccgc caacaattta tttaatccaa atatatgaa 16740
gtatattatt ccatagcctt tatttattta tatatttatt atataaaagc tttatttgtt 16800
ctaggttgtt catgaaatat ttttttggtt ttatctccgt tgtaagaaaa tcatgtgctt 16860
tgtgtcgcca ctcaactattg cagctttttc atgcattggt cagattgacg gttgattgta 16920
tttttgtttt ttatggtttt gtgttatgac ttaagtcttc atctctttat ctcttcatca 16980
ggtttgatgg ttaccttaata tggttcatgg gtacatgcat ggtaaatta ggtggccaac 17040
ttgtttgtga acgatagaat tttttttata ttaagtaaac tatttttata ttatgaaata 17100
ataataaaaa aaatatttta tcattattaa caaaatcata ttagttaatt tgtaactctt 17160
ataataaaag aaactctgta acattcactat tacatggtaa catctttcca ccctttcatt 17220
tgttttttgt ttgatgactt tttttcttgt ttaaatttat ttcccttctt ttaaatttgg 17280
aatacattat catcatatat aaactaaaat actaaaaaca ggattacaca aatgataaat 17340
aataacacaa atatttataa atctagctgc aatataatta aactagctat atcgatattg 17400
taaaataaaa ctgactgcat tgatactgat aaaaaaatat catgtgcttt ctggactgat 17460
gatgcagtat acttttgaca ttgcctttat tttatftttc agaaaagctt tcttagttct 17520
gggttcttca ttatttgttt cccatctcca ttgtgaattg aatcatttgc ttcgtgtcac 17580
aaatacaatt tagntaggta catgcattgg tcagattcac ggtttattat gtcatgactt 17640
aagttcatgg tagtacatta cctgccacgc atgcattata ttggttagat ttgataggca 17700
aatttggttg tcaacaatat aaatataaat aatgttttta tattacgaaa taacagtgat 17760
caaaacaac agttttatct ttattaacaa gatftttgtt ttgtttgatg acgtftttta 17820
atgtttacgc tttcccttct cttttgaatt tagaacctt tatcatcata aaatcaata 17880
ctaaaaaaat tacatatttc ataaataata acacaaatat ttttaaaaaa tctgaaataa 17940
taatgaacaa tattacatat tacacgaaa attcattaat aaaaatatta tataaataaa 18000
atgtaatagt agttatatgt aggaaaaaag tactgcacgc ataatatata caaaaagatt 18060
aaaatgaact attataaata ataacactaa attaatggtg aatcatatca aaataatgaa 18120

```

-continued

---

```

aaagtaaata aaatttgtaa ttaacttcta tatgtattac acacacaaat aataaataat 18180
agtaaaaaaa attatgataa atatttacca tctcataaga tatttaaaat aatgataaaa 18240
atatagatta ttttttatgc aactagctag ccaaaaagag aacacgggta tatataaaaa 18300
gagtaccttt aaattctact gtacttctct tattcctgac gtttttatat caagtggaca 18360
tacgtgaaga ttttaattat cagtctaaat atttcattag cacttaatac ttttctgttt 18420
tattcctatc ctataagtag tcccgattct cccaacattg cttattcaca caactaacta 18480
agaaagtctt ccatagcccc ccaagcggcc catggcctcc tccgaggacg tcatcaagga 18540
gttcatgcgc ttcaaggtgc gcatggaggg ctccgtgaac ggccacgagt tcgagatcga 18600
gggcgagggc gagggcgcgc cctacgaggg caccagacc gccaaagtga aggtgaccaa 18660
gggcggcccc ctgcccttcg cctgggacat cctgtcccc cagttccagt acggtcccaa 18720
ggtgtacgtg aagcaccocg ccgacatccc cgactacaag aagctgtcct tccccgaggg 18780
cttcaagtgg gagcgcgtga tgaacttcca ggacggcggc gtggtgaccg tgaccagga 18840
ctctccctg caggacggct ccttcatcta caaggtgaag ttcacggcg tgaacttccc 18900
ctccgacggc cccgtaatgc agaagaagac tatgggctgg gaggcctcca ccgagcgcct 18960
gtacccccgc gacggcgtgc tgaaggcga gatccacaag gccctgaagc tgaaggacgg 19020
cggccactac ctggtggagt tcaagtccat ctacatggcc aagaagcccg tgcagctgcc 19080
cggctactac tacgtggact ccaagctgga catcacctcc cacaacgagg actacacat 19140
cgtggagcag tacgagcgcg ccgagggcgc ccaccacctg ttcctgtagc ggccggccgc 19200
gacacaagtg tgagagtact aaataaatgc tttggttga cgaaatcatt aactaaata 19260
aaataatcaa agcttatata tgccttccgc taaggccgaa tgcaaagaaa ttggttcttt 19320
ctcgttatct tttgccactt ttactagtac gtattaatta ctacttaac atctttgttt 19380
acggctcatt atatccgtcg acgg 19404

```

```

<210> SEQ ID NO 62
<211> LENGTH: 30
<212> TYPE: DNA
<213> ORGANISM: Artificial sequence
<220> FEATURE:
<223> OTHER INFORMATION: Primer D6 fwd

```

```
<400> SEQUENCE: 62
```

```
gaattcggc cgcgatggct gctgctccca 30
```

```

<210> SEQ ID NO 63
<211> LENGTH: 30
<212> TYPE: DNA
<213> ORGANISM: Artificial sequence
<220> FEATURE:
<223> OTHER INFORMATION: Primer D6 rev

```

```
<400> SEQUENCE: 63
```

```
gaattcggc cgcgttactg cgccttacc 30
```

```

<210> SEQ ID NO 64
<211> LENGTH: 4322
<212> TYPE: DNA
<213> ORGANISM: Artificial sequence
<220> FEATURE:
<223> OTHER INFORMATION: Vector KS119

```

```
<400> SEQUENCE: 64
```

-continued

---

agcttttgat ccatgccctt catttgccgc ttattaatta atttggtaac agtccgtact	60
aatcagttac ttatocctcc cccatcataa ttaatcttgg tagtctcgaa tgccacaaca	120
ctgactagtc tcttggatca taagaaaaag ccaaggaaca aaagaagaca aaacacaatg	180
agagtatcct ttgcatagca atgtctaagt tcataaaatt caaacaaaaa cgcaatcaca	240
cacagtggac atcacttatc cactagctga tcaggatcgc cgcgtcaaga aaaaaaact	300
ggaccccaaa agccatgcac aacaacacgt actcacaag gtgtcaatcg agcagcccaa	360
aacattcaoc aactcaoccc atcatgagcc ctcacatttg ttgtttctaa cccaacctca	420
aactcgtatt ctcttccgcc acctcatttt tgtttatttc aacaccgcgc aaactgcatg	480
ccaccctcgt gccaatgtc catgcatggt aacaagacct atgactataa atagctgcaa	540
tctcggccca ggttttcac atcaagaacc agttcaatat cctagtcac cgtattaaag	600
aatttaagat atactgccc cgcaagatg aactaaaatg catgtaggtg taagagctca	660
tggagagcat ggaatattgt atccgacct gtaacagtat aataactgag ctccatctca	720
cttcttctat gaataaaca aggatgttat gatataataa cactctatct atgcacctta	780
ttgttctatg ataaatttcc tcttattatt ataaatcac tgaatcgtga cggcttatgg	840
aatgttcaa atagtacaaa aacaaatgtg tactataaga ctttctaacc aattctaacc	900
ttagcattgt gaacgagaca taagtgttaa gaagacataa caattataat ggaagaagtt	960
tgtctccatt tatatattat atattacca cttatgtatt atattaggat gttaaggaga	1020
cataacaatt ataagagag aagtttgat ccatttatat attatatact acccatttat	1080
atattatact tatccactta tttaatgtct ttataaggtt tgatccatga ttttctaat	1140
atthtagttg atagtgtat gaaagggtag tatttgaact ctcttactct gtataaagg	1200
tggatcatcc ttaaagtggt tctatttaatt tttattgctt cttacagata aaaaaaaaaat	1260
tatgagttgg tttgataaaa tattgaagga tttaaaataa taataaataa catataatat	1320
atgtatataa atttattata atataacatt tatctataaa aaagtaataa ttgtcataaa	1380
tctatacaat cgttttagct tctgggacga atctcaatta tttaaacgag agtaaacata	1440
tttgactttt tggttattta acaaatatt atttaacact atagaaatt tttttttta	1500
tcagcaaaga ataaatttaa attaagaagg acaatgggtg cccaatcctt atacaacca	1560
ctccacaag aagtcagt cagagacaac aaaaaaaca gcaaaggaaa ttttttaatt	1620
tgagttgtct tgtttgtgc ataatttatg cagtaaaaca ctacacataa cccttttagc	1680
agtagagcaa tggttgaccg tgtgcttagc ttcttttatt ttattttttt atcagcaaag	1740
aataaataaa ataaatgag acacttcagg gatgtttcaa caagcttggg tccgtcgacg	1800
gcgcgcccga tcatccggat atagttcctc ctttcagcaa aaaaccctc aagaccggt	1860
tagaggcccc aaggggttat gctagttatt gctcagcggg ggcagcagcc aactcagctt	1920
ccttcgggc tttgttagca gccggatcga tccaagctgt acctactat tctttgccc	1980
tcggacgagt gctggggcgt cggtttccac tatcggcgag tacttctaca cagccatcgg	2040
tccagacggc cgcgctctct cgggcgattt gtgtacgccc gacagtcccg gctccggatc	2100
ggacgattgc gtcgcatoga ccctgcgccc aagctgcac atcgaaattg ccgtcaacca	2160
agctctgata gagttgtca agaccaatgc ggagcatata cgcccggagc cgcggcgatc	2220
ctgcaagctc cggatgcctc cgtcgaagt agcgcgtctg ctgctccata caagccaacc	2280
acggcctcca gaagaagatg ttggcgacct cgtattggga atccccgaac atcgcctcgc	2340
tccagtcaat gaccgctggt atcgggcat tgtccgtcag gacattgttg gagccgaaat	2400

-continued

---

```

cgcgctgcac gaggtgcccg acttcggggc agtcctcggc ccaaagcacc agctcatcga 2460
gagcctgcgc gacggacgca ctgacgggtg cgtccatcac agtttgccag tgatacacat 2520
ggggatcagc aatcgccgat atgaaatcac gccatgtagt gtattgaccg attccttgcg 2580
gtccgaatgg gccgaaccgc ctgctctggc taagatcggc cgcagcgate gcacccatag 2640
cctccgcgac cggctgcaga acagcgggca gttcggtttc aggcaggtct tgcaacgtga 2700
caccctgtgc acggcgggag atgcaatagg tcaggctctc gctgaattcc ccaatgtcaa 2760
gcacttccgg aatcgggagc gcggccgatg caaagtgcgc ataaacataa cgatctttgt 2820
agaaaccatc ggccgagcta tttaccgca ggacatatcc acgcccctct acatcgaagc 2880
tgaaagcacg agattctctg cctcccgaga gctgcatcag gtcggagacg ctgtcgaact 2940
tttcgatcag aaacttctcg acagacgtcg cggtgagttc aggtttttcc atgggtatat 3000
ctccttctta aagttaaaca aaattatttc tagagggaaa ccggtgtggt ctccctatag 3060
tgagtcgat taatttcgcg ggatcgagat ctgatcaacc tgcattaatg aatcggccaa 3120
cgcgcgggga gagggggttt gcgtattggg cgctcttccg cttcctcgtc cactgactcg 3180
ctgcgctcgg tcgttcggct gcggcgagcg gtatcagctc actcaaaggc ggtaaacgg 3240
ttatccacag aatcagggga taacgcagga aagaacatgt gagcaaaagg ccagcaaaag 3300
gccaggaacc gtaaaaaggc cgcgttctg gcgtttttcc ataggctcgc cccccctgac 3360
gagcatcaca aaaatcgacg ctcaagtcag aggtggcgaa acccgacagg actataaaga 3420
taccaggcgt tccccctgg aagctccctc gtgcgctctc ctgttccgac cctgcccgtt 3480
accggatacc tgtccgcctt tctcccttcg ggaagcgtgg cgctttctca atgctcacgc 3540
tgtaggtatc tcagttcggg gtaggctggt cgctccaagc tgggctgtgt gcacgaacc 3600
cccgttcagc ccgaccgctg cgccttatcc ggtaactatc gtcttgagtc caaccggta 3660
agacacgact tatcgccact ggccagcagc actggttaaca ggattagcag agcagaggtat 3720
gtaggcggtg ctacagagtt cttgaagtgg tggcctaact acggctacac tagaaggaca 3780
gtatttggtt tctgcgctct gctgaagcca gttaccttcg gaaaaagagt tggtagctct 3840
tgatccggca aacaaccac cgctggtagc ggtggttttt ttgtttgcaa gcagcagatt 3900
acgcgcagaa aaaaaggatc tcaagaagat cctttgatct tttctacggg gtctgacgct 3960
cagtggaacg aaaactcacg ttaagggatt ttggtcatga cattaacctt taaaaatagg 4020
cgtatcacga ggcctcttcg tctcgcgctt ttcggtgatg acggtgaaaa cctctgacac 4080
atgcagctcc cggagacggt cacagcttgt ctgtaagcgg atgcggggag cagacaagcc 4140
cgtcagggcg cgtcagcggg tgttggcggg tgcgggggct ggcttaacta tgcggcatca 4200
gagcagattg tactgagagt gcaccatag gacatattgt cgttagaacg cggctacaat 4260
taatacataa ccttatgtat catacacata cgatttaggt gacactatag aacggcgcgc 4320
ca 4322

```

```

<210> SEQ ID NO 65
<211> LENGTH: 9420
<212> TYPE: DNA
<213> ORGANISM: Artificial sequence
<220> FEATURE:
<223> OTHER INFORMATION: Vector KS430
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: (2349)..(2349)
<223> OTHER INFORMATION: n is a, c, g, or t

```

-continued

&lt;400&gt; SEQUENCE: 65

ggccgcaagt atgaactaaa atgcatgtag gtgtaagagc tcatggagag catggaatat	60
tgtatccgac catgtaacag tataataact gagctccatc tcacttcttc tatgaataaa	120
caaaggatgt tatgatatat taacactcta tctatgcacc ttattgttct atgataaatt	180
tcctcttatt attataaatc atctgaatcg tgacggctta tggaatgctt caaatagtac	240
aaaaacaaat gtgtactata agactttcta aacaattcta accttagcat tgtgaacgag	300
acataagtgat taagaagaca taacaattat aatggaagaa gtttgtctcc atttatatat	360
tatatattac ccacttatgt attatattag gatgttaagg agacataaca attataaaga	420
gagaagtttg tatecattta tatattatat actaccatt tatatattat acttatccac	480
ttatttaatg tctttataag gtttgatcca tgatatttct aatattttag ttgatatgta	540
tatgaaaggg tactatttga actcctttac tctgtataaa ggttggatca tccttaaagt	600
gggtctattt aattttattg cttcttacag ataaaaaaaa aattatgagt tggtttgata	660
aaatattgaa ggatttaaaa taataataaa taacatataa tatatgtata taaatttatt	720
ataatataac atttatctat aaaaaagtaa atattgtcat aaatctatac aatcgtttag	780
ccttgctgga cgaatctcaa ttatttaaac gagagtaaac atatttgact ttttggttat	840
ttaacaaatt attatttaac actatatgaa attttttttt ttatcagcaa agaataaaat	900
taaatgaaga aggacaatgg tgtcccaatc cttatacaac caacttccac aagaaagtca	960
agtcagagac aacaaaaaaaa caagcaagg aaatttttta atttgagttg tcttgtttgc	1020
tgcataatth atgcagtaaa aactacaca taaccctttt agcagtagag caatggttga	1080
ccgtgtgctt agcttctttt attttatttt ttatcagca aagaataaat aaaaaaaat	1140
gagacacttc agggatgtht caacaagctt ggatcctcga agagaagggg taataacaca	1200
cttttttaac atttttaaca caaatttttag ttatttaaaa atttattaaa aaatttaaaa	1260
taagaagagg aactctttaa ataaatctaa cttacaaaa ttatgatttt taataagttt	1320
tcaccaataa aaaatgtcat aaaaatattg taaaaagtat attatcaata ttctctttat	1380
gataaataaa aagaaaaaaaa aaataaaagt taagtgaaaa tgagattgaa gtgactttag	1440
gtgtgtataa atatatcaac cccgccaaaca atttatttaa tccaaatata ttgaagtata	1500
ttattccata gcctttatth atttatatat ttattatata aaagctttat ttgttctagg	1560
ttgttcatga aatatttttt tggttttatc tccgttgtaa gaaaatcatg tgctttgtgt	1620
cgccactcac tattgcagct ttttcatgca ttggtcagat tgacggttga ttgtattttt	1680
gttttttatg gttttgtgth atgacttaag tcttcatctc tttatctctt catcaggtht	1740
gatggttacc taatattgth catgggtaca tgcattgtht aattaggtgg ccaactttgt	1800
tgtgaacgat agaatttttt ttatattaag taaactatth ttatattatg aaataataat	1860
aaaaaaaaata ttttatcatt attaacaaaa tcatattatg taatttgtht actctataat	1920
aaaagaaata ctgtaacatt cacattacat ggtaacatct tccaccctt tcatttgtht	1980
ttgtttgat gacttttttt cttgtttaaa tttatttccc tctttttaa tttggaatac	2040
attatcatca tatataaact aaaatactaa aaacaggatt acacaaatga taaataataa	2100
cacaaatatt tataaatcta gctgcaatat atttaaacta gctatatcga tattgtaaaa	2160
taaaactagc tgcattgata ctgataaaaa aatattcatg gctttctgga ctgatgatgc	2220
agtatactth tgacattgth tttattttat ttttcagaaa agctttctta gttctggtht	2280
cttcattatt tgtttcccat ctccattgtg aattgaatca tttgcttctg gtcacaaata	2340

-continued

---

caatttagnt aggtacatgc attggtcaga ttcacggttt attatgcat gacttaagtt	2400
catggtagta cattacctgc cacgcatgca ttatattggt tagatttgat aggcaaattt	2460
ggttgcaac aatataaata taaataatgt ttttatatta cgaataaaca gtgatcaaaa	2520
caaacagttt tatctttatt aacaagattt tgtttttggt tgatgacggt ttttaagtt	2580
tacgctttcc cccttctttt gaatttagaa cactttatca tcataaaatc aaatactaaa	2640
aaaattacat atttcataaa taataacaca aatattttta aaaaatctga aataaataatg	2700
aacaatatta catattatca cgaaaattca ttaataaaaa tattatataa ataaaatgta	2760
atagtagtta tatgtaggaa aaaagtactg cacgcataat atatacaaaa agattaaat	2820
gaactattat aaataataac actaaattaa tggatgaatca tatcaaaaata atgaaaaagt	2880
aaataaaaatt tgtaattaac ttctatatgt attacacaca caaataataa ataatagtaa	2940
aaaaaattat gataaatatt taccatctca taagatattt aaaataatga taaaaatata	3000
gattattttt tatgcaacta gctagccaaa aagagaacac gggatatatat aaaaagagta	3060
cctttaaatt ctactgtact tcttttattc ctgacgtttt tatatcaagt ggacatacgt	3120
gaagatttta attatcagtc taaatatttc attagcactt aatacttttc tgttttattc	3180
ctatcctata agtagtcccg attctcccaa cattgcttat tcacacaact aactaagaaa	3240
gtcttccata gcccccaag cggcccatgg cctctccga ggacgtcacc aaggagtcca	3300
tgcgcttcaa ggtgcgcatg gagggctccg tgaacggcca cgagtccgag atcgagggcg	3360
agggcgaggg cgcgccctac gagggcacc cagaccgcaa gctgaagggt accaagggcg	3420
gccccctgcc ctctgcctgg gacatcctgt cccccagtt ccagtacggc tccaagggtg	3480
acgtgaagca ccccgccgac atccccgact acaagaagct gtccttcccc gagggcttca	3540
agtgggagcg cgtgatgaac ttcgaggacg gcggcgtggt gaccgtgacc caggactcct	3600
ccctgcagga cggtctcttc atctacaagg tgaagtccat cggcgtgaac ttccccctcg	3660
acggccccgt aatgcagaag aagactatgg gctgggaggc ctccaccgag cgctgtacc	3720
cccgcgacgg cgtgctgaag ggcgagatcc acaaggccct gaagctgaag gacggcggcc	3780
actacctggt ggagtccaag tccatctaca tggccaagaa gcccggtgcag ctgccccggt	3840
actactacgt ggactccaag ctggacatca cctcccacaa cgaggactac accatcgtgg	3900
agcagtacga gcgcgcccag ggcgcgccacc acctgttctc gtagcggccc gccgcgacac	3960
aagtgtaga gtactaaata aatgctttgg ttgtacgaaa tcattacact aaataaata	4020
atcaaagctt atatatgctt tccgctaagg ccgaatgcaa agaaattggt tctttctcgt	4080
tatcttttgc cacttttact agtacgtatt aattactact taatcatctt tgtttacggc	4140
tcattatata cgtcgacggc gcgcgctct agagggccca attcgccta tagtgagtcg	4200
tattacaatt cactggccgt cgttttacia cgtcgtgact gggaaaacc ttggcgttacc	4260
caacttaac gccttgacgc acatccccct ttcgccagct ggcgtaatag cgaagaggcc	4320
cgaccgac gccttccca acagttgcgc agcctatacg tacggcagtt taaggtttac	4380
acctataaaa gagagagccg ttatcgtctg tttgtggatg tacagagtga tattattgac	4440
acgccggggc gacggatggt gatccccctg gccagtgcac gtctgctgtc agataaagtc	4500
tcccgtgaac tttaccgggt ggtgcatatc ggggatgaaa gctggcgcac gatgaccacc	4560
gatatggcca gtgtgccggt ctccgttatc ggggaagaag tggctgatct cagccaccgc	4620
gaaaatgaca tcaaaaacgc cattaacctg atgttctggg gaatataaat gtcaggcatg	4680

-continued

---

agattatcaa aaaggatcctt cacctagatc cttttcacgt agaaagccag tccgcagaaa	4740
cggtgctgac cccggatgaa tgtcagctac tgggctatct ggacaaggga aaacgcaagc	4800
gcaaagagaa agcaggtagc ttgcagtggg cttacatggc gatagctaga ctgggcggtt	4860
ttatggacag caagcgaacc ggaattgcca gctggggcgc cctctggtaa ggttgggaag	4920
cctgcaaag taaactggat ggctttcttg ccgccaagga tctgatggcg caggggatca	4980
agctctgac aagagacagg atgaggatcg tttcgcata ttgaacaaga tggattgac	5040
gcaggttctc cggccgcttg ggtggagagg ctattcggct atgactgggc acaacagaca	5100
atcggctgct ctgatgcgc cgtgttcccg ctgtcagcgc aggggcgccc ggttcttttt	5160
gtcaagaccg acctgtccgg tgcctgaat gaactgcaag acgaggcagc gcggctatcg	5220
tggctggcca cgacggcgt tccttgcca gctgtgctcg acgttgcac tgaagcggga	5280
agggactggc tgctattggg cgaagtgcg gggcaggatc tcctgtcacc tcacctgct	5340
cctgccgaga aagtatccat catggctgat gcaatgcggc ggctgcatac gcttgatccg	5400
gctacctgcc cattcgacca ccaagcgaat catcgcacg agcgagcacg tactcggatg	5460
gaagccggtc ttgtcgatca ggatgatctg gacgaagagc atcaggggct cgcgccagcc	5520
gaactgttcg ccaggctcaa ggcgagcatg cccgacggcg aggatctcgt cgtgacctat	5580
ggcgatgctt gcttgccgaa tatcatggtg gaaaatggcc gctttctgg attcatcgac	5640
tgtggccggc tgggtgtggc ggaccgctat caggacatag cgttggtacc ccgtgatatt	5700
gctgaagagc ttggcggcga atgggctgac cgcttctcgt tgetttacgg tatcgccgct	5760
cccgatctgc agcgcacgc cttctatcgc cttcttgacg agttctctg aattattaac	5820
gcttacaatt tcctgatgcg gtattttctc cttacgcacc tgtgcggtat ttcacaccgc	5880
atcaggtggc acttttcggg gaaatgtgcg cggaaaccct atttgttat ttttctaaat	5940
acattcaaat atgtatccgc tcatgagaca ataaccctga taaatgcttc aataatagca	6000
cgtgaggagg gccaccatgg ccaagttagc cagtgcctt ccggtgctca ccgcgcgcga	6060
cgtgcccgga gcggctgagt tctggaccga ccggctcggg ttctcccggg acttcgtgga	6120
ggacgacttc gccgggtgtg tccgggacga cgtgaccctg ttcacagcg cggctccagga	6180
ccagtggtg cgcgacaaca cctggcctg ggtgtgggtg cgcggcctgg acgagctgta	6240
cgcgagtggt tcggaggtcg tgtccacgaa cttccgggac gcctccgggc cggccatgac	6300
cgagatcggc gagcagcgt gggggcggga gttccgcctg cgcgaccggc ccggcaactg	6360
cgtgcacttc gtggccgagg agcaggactg acacgtgcta aaacttcatt ttaatttaa	6420
aaggatctag gtgaagatcc tttttgataa tctcatgacc aaaatccctt aacgtgagtt	6480
ttcgttccac tgagcgtcag accccgtaga aaagatcaaa ggatcttctt gagatccttt	6540
ttttctgcgc gtaatctgct gcttgcaaac aaaaaacca ccgctaccag cggtggtttg	6600
tttgccggat caagagctac caactctttt tccgaaggta actggcttca gcagagcgca	6660
gataccaaat actgttcttc tagttagacc gtagttaggc caccacttca agaactctgt	6720
agcaccgcct acatacctcg ctctgctaact cctgttacca gtgctgctg ccagtgccga	6780
taagtcgtgt cttaccgggt tggactcaag acgatagtta ccggataagg cgcagcggtc	6840
gggctgaacg gggggtcgt gcacacagcc cagcttgag cgaacgacct acaccgaact	6900
gagataccta cagcgtgagc tatgagaaaag cgcacgcctt cccgaaggga gaaaggcggga	6960
caggtatccg gtaagcggca gggctcggaac aggagagcgc acgagggagc ttccaggggg	7020
aaacgcctgg tatctttata gtctgtcgg gtttcgccac ctctgacttg agcgtcgatt	7080

-continued

---

tttgtgatgc	tcgtcagggg	ggcggagcct	atggaaaaac	gccagcaacg	cggccttttt	7140
acggttctctg	gccttttget	ggccttttgc	tcacatgttc	tttctctcgt	tatcccctga	7200
ttctgtggat	aaccgtatta	ccgcctttga	gtgagctgat	accgctcgcc	gcagccgaac	7260
gaccgagcgc	agcagtcag	tgagcgagga	agcggaaagag	cgcccaatac	gcaaaccgcc	7320
tctcccccg	cgttggccga	ttcattaatg	cagctggcac	gacaggtttc	ccgactggaa	7380
agcgggcagt	gagcgcaacg	caattaatgt	gagttagctc	actcattagg	caccccaggc	7440
tttacacttt	atgcttcogg	ctcgtatgtt	gtgtggaatt	gtgagcggat	aacaatttca	7500
cacaggaaac	agctatgacc	atgattacgc	caagctattt	aggtgacgcg	ttagaatact	7560
caagctatgc	atcaagcttg	gtaccgagct	cggatccact	agtaaccgcc	gccagtgtgc	7620
tggaattcag	gggcgcgccc	caatttatta	tcctctattt	cctgacattt	taatccatcc	7680
acctatgtca	aaaacttata	gaaaatgtca	acttccaaac	aaaacataat	tgaacttcgc	7740
aaataaattc	ttaataatat	taaaaaatgt	tacttaatta	tttcttcaac	cccattttcc	7800
gcgctagcgc	cggacaaaga	ctctagttaa	atatagaagt	ttccgattct	catcgtataa	7860
aacggtgact	ttggcgggct	ttcatgtgta	acaaattggt	ttaacaaacc	actgcctagt	7920
cgttttagtgt	agaatcagcg	catggaactc	cgattggagc	gtgactttca	cgtgcgggag	7980
gcccaccacc	acagcggggc	ttacgctcta	agaatctcgc	ccacggtttt	cttcatctcc	8040
cccccgccaa	gtgtctccct	cgttcgccac	ttctcatcat	gttacaggga	ccataaaaat	8100
ggcgtatttc	ttcagccccg	ggtataaata	cacacatgat	cctgtggtgg	gttcttcac	8160
aagttacatc	tccttctggt	ttttgtattg	caagtgtttg	tattttttgc	ctccgagaga	8220
aaatcgcggc	cgcattggaga	gatctcaacg	gcagtctcct	ccgccaccgt	cgccgtctct	8280
ctcctcgtcc	tcctctctcc	cggacaccgt	cctcgtccct	cccgaaaga	ggcggagggc	8340
ggcgacggcc	aagcccgggc	ccgagcctaa	taagaggatc	cgcaaggacc	ccgccgccgc	8400
cgcccggggg	aagaggagct	ccgtctacag	gggagtcacc	aggcacaggc	ggaaggcag	8460
gttcgaggcg	catctctggg	acaagcactg	cctcgcgcgc	ctccacaaca	agaagaaagg	8520
caggcaagtc	tacctggggg	cgtatgacag	cgaggaggca	gctgctcgtg	cctatgacct	8580
cgcagctctc	aagtactggg	gtcctgagac	tctgctcaac	ttccctgtgg	aggattactc	8640
cagcgagatg	ccggagatgg	aggccgtgtc	ccgggaggag	tacctggcct	ccctccgccg	8700
caggagcagc	ggcttctcca	ggggcgtctc	caagtacaga	ggcgtcgcca	ggcatcacca	8760
caacgggagg	tgggaggcac	ggattgggcg	agtctttggg	aacaagtacc	tctacttggg	8820
aacatttgac	actcaagaag	aggcagccaa	ggcctatgac	cttgccggcca	ttgaataaccg	8880
tggcgtcaat	gctgtaacca	acttcgacat	cagctgctac	ctggaccacc	cgtgttccct	8940
ggcacagctc	caacaggagc	cacaggtggt	gccggcactc	aaccaagaac	ctcaacctga	9000
tcagagcgaa	accggaacta	cagagcaaga	gccggagtca	agcgaagcca	agacaccgga	9060
tggcagtgca	gaacccgatg	agaacgcggt	gcctgacgac	accgcgagc	ccctcaccac	9120
agtcagcagc	agcatcgaag	aggccttctg	gagcccttgc	atggattacg	agctagacac	9180
catgtcgaga	ccaaaacttg	gcagctcaat	caatctgagc	gagtggttcg	ctgacgcaga	9240
cttcgactgc	aacatcggat	gcctgttcga	tgggtgttct	gcgctgacg	aaggaagcaa	9300
ggatgggtga	ggtctggcag	atctcagctc	gtttgaggca	ggtgatgtcc	agctgaagga	9360
tgttctttcg	gatatggaag	aggggataca	acctccagcg	atgatcagtg	tgtgcaacgc	9420

-continued

---

```

<210> SEQ ID NO 66
<211> LENGTH: 19215
<212> TYPE: DNA
<213> ORGANISM: Artificial sequence
<220> FEATURE:
<223> OTHER INFORMATION: Vector ARAL078
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: (17405)..(17405)
<223> OTHER INFORMATION: n i s a, c, g, o r t

<400> SEQUENCE: 66

cgcgccctcga gtggggcgat cccccgggct gcaggaattc actggccgctc gttttacaac    60
gtcgtgactg ggaaaacctt ggcgttaccc aacttaatcg ccttgacgca catccccctt    120
tcgccagctg gcgtaatagc gaagaggccc gcaccgatcg cccttcccaa cagttgcgca    180
gcctgaatgg cgaatggatc gatccatcgc gatgtacctt ttgttagtca gcctctcgat    240
tgctcatcgt cattacacag taccgaagtt tgatcgatct agtaacatag atgacaccgc    300
gcgcgataat ttatcctagt ttgcgcgcta tttttgttt tctatcgcgt attaaatgta    360
taattgctgg actctaatca taaaaacca tctcataaat aacgtcatgc attacatggt    420
aattattaca tgcttaacgt aattcaacag aaattatatg ataacatcg caagaccggc    480
aacaggattc aatcttaaga aactttattg ccaaatgttt gaacgatctg cttcgacgca    540
ctccttcttt actccaccat ctgcctctta ttgaaaacgt gggtagcacc aaaacgaatc    600
aagtcgctgg aactgaagtt accaatcacg ctggatgatt tgccagtggg attaactctg    660
cctttccccg catgaataat attgatgaat gcatgogtga ggggtagttc gatggtggca    720
atagctgcaa ttgccgcgac atcctccaac gagcataatt cttcagaaaa atagcgatgt    780
tccatgttgt cagggcatgc atgatgcacg ttatgaggtg acggtgctag gcagatttcc    840
ctcaaagttt catagtcagt atcatattca tcattgcatt cctgcaagag agaattgaga    900
cgcaatccac acgctgcggc aaccttccgg cgttcgtggt ctatttgctc ttggacgttg    960
caaacgtaag tgttggatcg atccgggggtg ggcgaagaac tccagcatga gatccccgcg    1020
ctggaggatc atccagccgg cgtcccggaa aacgattccg aagcccaacc tttcatagaa    1080
ggcggcgggt gaatcgaat ctctgatggc caggttgggc gtcgcttggg cggtcatttc    1140
gaaccccaga gtcccgcctc gaagaactcg tcaagaaggg gatagaaggg gatgctctgc    1200
gaatcgggag cggcgatacc gtaaagcacg aggaagcggg cagccattc gccgccaagc    1260
tcttcagcaa taccacgggt agccaacgct atgtctgat agcggtcctc cacaccacgc    1320
cggccacagt cgtgaatcc agaaaagcgg ccattttcca ccatgatatt cggcaagcag    1380
gcatcgccat gggtcacgac gagatcctcg ccgtcgggca tgcgcgcctt gagcctggcg    1440
aacagttcgg ctggcgcgag ccctgatgc tcttcgtcca gatcatcctg atcgacaaga    1500
ccggttcca tccagtagc tgctcgctcg atgcatggtt tcgcttgggt gtcgaatggg    1560
caggtagcgg gatcaagcgt atgcagccgc cgcattgcat cagccatgat ggatactttc    1620
tcggcaggag caaggtgaga tgacaggaga tcctgccccg gcaacttcgcc caatagcagc    1680
cagtcctctc ccgcttcagt gacaacgctg agcacagctg cgcaaggaac gcccgctcgtg    1740
gccagccacg atagccgcgc tgccctctcc tgcagttcat tcagggcacc ggacaggctc    1800
gtcttgacaa aaagaaccgg gcgcccctgc gctgacagcc ggaacacggc ggcacagag    1860
cagccgattg tctgttgtgc ccagtcatag ccgaatagcc tctccacca agcggccgga    1920
gaacctgcgt gcaatccatc ttgttcaatc atgcgaaaac atccccgcaa gcttgagac    1980

```

-continued

---

tggtgatttc	agcgtgtcct	ctccaaatga	aatgaacttc	cttatataga	ggaaggtct	2040
tgcaaggat	agtgggattg	tgcgtcatcc	cttacgtcag	tggagatata	acatcaatcc	2100
acttgctttg	aagacgtggg	tggaaactct	tctttttcca	cgatgctcct	cgtgggtggg	2160
ggtccatctt	tgggaccact	gtcggcagag	gcattctcaa	cgatggcctt	tcctttatcg	2220
caatgatggc	attttagtag	gccaccttcc	ttttccacta	tcttcacaat	aaagtgcag	2280
atagctgggc	aatggaatcc	gaggagggtt	cggatatta	ccctttgttg	aaaagtctca	2340
attgcccttt	ggtcttctga	gactgtatct	ttgatatttt	tggagtagac	aagcgtgtcg	2400
tgctccacca	tgttgacgaa	gattttcttc	ttgtcattga	gtcgtaaagag	actctgtatg	2460
aactgttcgc	cagtccttac	ggcgagttct	gttaggtcct	ctatttgaat	ctttgactcc	2520
atggcctttg	attcagtggt	aactaccttt	ttagagactc	caatctctat	tacttgctct	2580
ggtttgtgaa	gcaagccttg	aatcgtccat	actggaatag	tacttctgat	cttgagaaat	2640
atatctttct	ctgtgttctt	gatgcagtta	gtcctgaatc	ttttgactgc	atctttaacc	2700
ttcttgggaa	ggtatttgat	ctcctggaga	ttattgctcg	ggtagatcgt	cttgatgaga	2760
cctgtcgcgt	aagcctctct	aaccatctgt	gggttagcat	tctttctgaa	attgaaaagg	2820
ctaactctct	cattatcagt	ggtgaacatg	gtatcgtcac	cttctccgtc	gaacttctcg	2880
actagatcgt	agagatagag	gaagtcgtcc	attgtgatct	ctggggcaaa	ggagatctga	2940
attaattcga	tatggtggat	ttatcacaaa	tgggaccgcg	cgccgacaga	ggtgtgatgt	3000
taggccagga	ctttgaaaat	ttgcgcaact	atcgtatagt	ggccgacaaa	ttgacgccga	3060
gttgacagac	tgcttagcat	ttgagtgaat	tatgtgaggt	aatgggctac	actgaattgg	3120
tagctcaaac	tgtcagtatt	tatgtatatg	agtgtatatt	ttcgcataat	ctcagaccaa	3180
tctgaagatg	aaatgggtat	ctgggaatgg	cgaaatcaag	gcatacgatcg	tgaagtttct	3240
catctaagcc	cccatttggg	cgtgaatgta	gacacgtcga	aataaagatt	tcogaattag	3300
aataatttgt	ttattgcttt	cgctataaaa	tacgacggat	cgtaatttgt	cgttttatca	3360
aatgtactt	tcattttata	ataacgctgc	ggacatctac	atttttgaat	tgaaaaaaaa	3420
ttggtaatta	ctctttcttt	ttctccatat	tgaccatcat	actcattgct	gatccatgta	3480
gatttcccgg	acatgaagcc	atttacaatt	gaatataatc	tgcgcccgt	gccgctttgc	3540
acccggtgga	gcttgcattg	tggtttctac	gcagaactga	gccggttagg	cagataattt	3600
ccattgagaa	ctgagccatg	tgcaccttcc	ccccaacacg	gtgagcgacg	gggcaacgga	3660
gtgatccaca	tgggactttt	aaacatcctc	cgtcggatgg	cgttgcgaga	gaagcagtcg	3720
atccgtgaga	tcagccgacg	caccgggcag	gcgcgcaaca	cgatcgcaaa	gtatttgaac	3780
gcaggtacaa	tcgagccgac	gttcacgcgg	aacgaccaag	caagctagct	ttaatgcggt	3840
agtttatcac	agttaaattg	ctaacgcagt	caggcaccgt	gtatgaaatc	taacaatgcg	3900
ctcatcgtea	tcctcggcac	cgtcaccctg	gatgctgtag	gcataggctt	ggttatgccg	3960
gtactgcccg	gcctcttgcg	ggatatactc	cattccgaca	gcatagccag	tcactatggc	4020
gtgctgctag	cgctatatgc	ggtgatgcaa	tttctatgcy	caccogttct	cggagcactg	4080
tccgaccgct	ttggccgccc	cccagtcctg	ctcgtctcgc	tacttggagc	cactatcgac	4140
tacgcgatca	tggcgaccac	accctcctg	tggccaacc	cctccgctgc	tatagtgcag	4200
tcggcttctg	acgttcagt	cagccgtctt	ctgaaaacga	catgtccgac	aagtcctaag	4260
ttacgcgaca	ggctgcccgc	ctgccctttt	cctggcgttt	tcttgctcgc	tgttttagtc	4320

-continued

---

gcataaagta gaataacttgc gactagaacc ggagacatta cgccatgaac aagagcgccg	4380
ccgctggcct gctgggtat gcccgcgta gcaccgacga ccaggacttg accaaccaac	4440
gggcccgaact gcaecgggcc ggtgacacca agctgttttc cgagaagatc accggcacca	4500
ggcgcgaccg cccggagctg gccaggatgc ttgaccacct acgcccctggc gacgttgtga	4560
cagtgaccag gctagaccgc ctggcccgcga gcaccgcga cctactggac attgcccagc	4620
gcatccagga ggccggcgcg ggccctgctg gctggcaga gccgtgggcc gacaccacca	4680
cgccggccgg ccgcatggtg ttgaccgtgt tcgcccgcac tgccgagttc gagcgttccc	4740
taatcatcga ccgcccggcg agcgggcccg aggcggccaa ggcccagggc gtgaagtttg	4800
gccccgccc taccctcacc ccggcacaga tcgcccgcgc ccgcgagctg atcgaccag	4860
aaggccgcac cgtgaaagag gcggtgcac tgcttggcgt gcatcgctcg accctgtacc	4920
gcgcaactga gcgcagcgag gaagtgcgc ccaccgaggc caggcggcgc ggtgccttc	4980
gtgaggacgc attgaccgag gccgacgcc tgggggccgc cgagaatgaa cgccaagagg	5040
aacaagcatg aaaccgcacc aggacggcca ggacgaaccg tttttcatta ccgaagagat	5100
cgaggcggag atgatcggcg ccgggtacgt gttcgagccg cccgcgcacg tctcaaccgt	5160
gcggtcgcac gaaatcctgg ccggtttgtc tgatgccaag ctggcggcct ggccggccag	5220
cttgccgct gaagaaaccg agcggcggcg tctaaaaagg tgatgtgtat ttgagtaaaa	5280
cagcttgcgt catcggtcgc ctgcgtatat gatcgatga gtaataaac aaatacgcga	5340
gggaacgcac gaagtatcgc ctgtacttaa ccagaaaggc gggtcaggca agacgaccat	5400
cgcaaccat ctagcccgcg ccctgcaact cgccggggcc gatgttctgt tagtcgattc	5460
cgatccccag ggcagtgccc gcgattgggc ggcctgcccg gaagatcaac cgctaaccgt	5520
tgccggcacc gaccgcccga cgattgaccg cgacgtgaag gccatcggcc ggcgcgactt	5580
cgtagtgatc gacggagcgc cccaggcggc ggacttggct gtgtcccga tcaaggcagc	5640
cgacttcgtg ctgattccgg tcagcccag cccttacgac atatgggcca ccgcccacct	5700
ggtggagctg gttaaagcgc gcattgaggt cacggatgga aggctacaag cggcctttgt	5760
cgtgtcggcg gcgatcaaag gcacgcgcac cggcggtagg gttgccgagg cgctggccgg	5820
gtacgagctg cccattcttg agtcccgtat cacgcagcgc gtgagctacc caggcactgc	5880
cgcccccggc acaaccgttc ttgaatcaga acccgagggc gacgctgccc gcgaggtcca	5940
ggcgctggcc gctgaaatta aatcaaaact catttgagtt aatgaggtaa agagaaaatg	6000
agcaaaagca caaacacgct aagtgcggcg cgtccgagcg cacgcagcag caaggctgca	6060
acgttggcca gcctggcaga cacgccagcc atgaagcggg tcaactttca gttgccggcg	6120
gaggatcaca ccaagctgaa gatgtacgcg gtacgccaag gcaagaccat taccgagctg	6180
ctatctgaat acatcgcgca gctaccagag taaatgagca aatgaataaa tgagtagatg	6240
aattttagcg gctaaaggag gcgcatgga aatcaagaa caaccaggca ccgacgccgt	6300
ggaatgcccc atgtgtggag gaacggggcg ttggccaggc gtaagcggct gggttgtctg	6360
ccggccctgc aatggcactg gaacccccaa gcccgaggaa tcgcgctgag cggtccgaaa	6420
ccatccggcc cggtaacaat cggcggcgcg ctgggtgatg acctggtgga gaagttgaa	6480
gcccgcgagg ccgcccagcg gcaacgcac gaggcagaag cacgcccggc tgaatcgtgg	6540
caagcggcgg ctgategaat ccgcaagaa tcccggcaac cgcccgcagc cggtgccggc	6600
tcgattagga agccgcccga gggcgacgag caaccagatt ttttcgttcc gatgctctat	6660
gacgtgggca ccccgatag tcgcagcacc atggaagcgg ccgttttccg tctgtcgaag	6720

-continued

---

cgtgaccgac	gagctggcga	ggtgatccgc	tacgagcttc	cagacgggca	cgtagaggtt	6780
tccgcagggc	cggccggcat	ggccagtgtg	tgggattacg	acctggact	gatggcggtt	6840
tcccatctaa	ccgaatccat	gaaccgatac	cgggaagggg	agggagacaa	gcccggccgc	6900
gtgttccgtc	cacacgttgc	ggacgtactc	aagttctgcc	ggcgagccga	tggcgaaaag	6960
cagaaagacg	acctggtaga	aacctgcatt	cggttaaaca	ccacgcacgt	tgccatgcag	7020
cgtacgaaga	aggccaagaa	cggcccgcctg	gtgacgggat	ccgaggggta	agccttgatt	7080
agccgctaca	agatcgtaaa	gagcgaacc	ggcgccggcg	agtacatcga	gatcgagcta	7140
gctgattgga	tgtaccgcga	gatcacagaa	ggcaagaacc	cggacgtgct	gacggttcac	7200
cccgattact	ttttgatcga	tcccggcatc	ggccgttttc	tctaccgcct	ggcacgccgc	7260
gccgcaggca	aggcagaagc	cagatggttg	ttcaagacga	tctacgaacg	cagtggcagc	7320
gccggagagt	tcaagaagtt	ctgtttcacc	gtgcgcaagc	tgatcgggtc	aaatgacctg	7380
ccggagtacg	atttgaagga	ggaggcgggg	caggctggcc	cgatcctagt	catgcgctac	7440
cgcaacctga	tcgagggcga	agcatccgcc	ggttcctaata	gtacggagca	gatgctaggg	7500
caaattgcc	tagcagggga	aaaaggtcga	aaaggtctct	ttctgtgga	tagcacgtac	7560
attgggaacc	caaagccgta	cattgggaac	cggaacccgt	acattgggaa	cccaaagccg	7620
tacattggga	accggtcaca	catgtaagtg	actgatataa	aagagaaaaa	aggcgatttt	7680
tccgcctaaa	actctttaa	acttattaaa	actcttaaaa	cccgcctggc	ctgtgcataa	7740
ctgtctggcc	agcgcacagc	cgaagagctg	caaaaagcgc	ctacccttcg	gtcgctgcgc	7800
tccctacgcc	ccgccccttc	gcgtcggcct	atcgcggccg	ctggccgctc	aaaaatgget	7860
ggcctacggc	caggcaatct	accagggcgc	ggacaagccg	cgccgctgcc	actcgaccgc	7920
cggcgcccac	atcaaggcac	cctgcctcgc	gcgtttcggg	gatgacgggtg	aaaaacctctg	7980
acacatgcag	ctcccggaga	cggtcacagc	ttgtctgtaa	gcggatgccg	ggagcagaca	8040
agcccgtcag	ggcgcgtcag	cgggtgttgg	cgggtgtcgg	ggcgcagcca	tgaccagtc	8100
acgtagcgat	agcggagtgt	atactggctt	aactatgcgg	catcagagca	gattgtactg	8160
agagtgcacc	atatgcggtg	tgaaataccg	cacagatcgc	taaggagaaa	ataccgcac	8220
aggcgctctt	ccgcttcctc	gctcactgac	tcgctgcgct	cggtcgttcg	gctgcggcga	8280
gcggtatcag	ctcactcaaa	ggcggtaata	cggttatcca	cagaatcagg	ggataacgca	8340
ggaaagaaca	tgtgagcaaa	aggccagcaa	aaggccagga	accgtaaaaa	ggccgcggtg	8400
ctggcgtttt	tccatagget	ccgccccct	gacgagcacc	acaaaaatcg	acgctcaagt	8460
cagaggtggc	gaaacccgac	aggactataa	agataaccag	cgtttcccc	tggaaagctcc	8520
ctcgtgcgct	ctcctgttcc	gacctgcgcg	cttaccggat	acctgtccgc	ctttctccct	8580
tccggaagcg	tggcgcttcc	tcatagetca	cgctgtaggt	atctcagttc	ggtgtaggtc	8640
gttcgctcca	agctgggctg	tgtgcacgaa	cccccgcttc	agcccagccg	ctgcgcctta	8700
tccgtaact	atcgtcttga	gtccaacccg	gtaagacacg	acttatcgcc	actggcagca	8760
gccactggta	acaggattag	cagagcgagg	tatgtaggcg	gtgctacaga	gttcttgaag	8820
tgggtggccta	actacggcta	cactagaagg	acagtatttg	gtatctgcgc	tctgctgaag	8880
ccagttacct	tcggaaaaag	agttggtagc	tcttgatccg	gcaaaaaaac	caccgctggt	8940
agcggtggtt	tttttggttg	caagcagcag	attacgcgca	gaaaaaaagg	atctcaagaa	9000
gatcctttga	tcttttctac	ggggctctgac	gctcagtgga	acgaaaaactc	acgttaaggg	9060

-continued

---

atthttgtca	tgagattatc	aaaaaggatc	ttcacctaga	tccttttaaa	ttaaaaatga	9120
agthtttaaat	caatctaaag	tatatatgag	taaacttggt	ctgacagtta	ccaatgctta	9180
atcagtgagg	cacctatctc	agcgatctgt	ctatttcggt	catccatagt	tgctgactc	9240
cccgtcgtgt	agataactac	gatacgggag	ggcttaccat	ctggccccag	tgctgcaatg	9300
ataccgcgag	accacgcctc	accggctcca	gatttatcag	caataaacca	gccagccgga	9360
agggccgagc	gcagaagtgg	tctgcaact	ttatccgctc	ccatccagtc	tattaattgt	9420
tgccgggaag	ctagagtaag	tagttcgcca	gttaaatagt	tgcgcaacgt	tgttgccatt	9480
gctacaggca	tcgtggtgtc	acgctcgctg	tttggtatgg	cttcattcag	ctccggttcc	9540
caacgatecaa	ggcgagttac	atgatcccc	atggttgca	aaaaagcggg	tagctccttc	9600
ggctctccga	tcgttgcag	aagtaagttg	gccgcagtg	tatcactcat	ggttatggca	9660
gcactgcata	attctcttac	tgctcatgcca	tccgtaagat	gcttttctgt	gactggtgag	9720
tactcaacca	agtcattctg	agaatagtgt	atgcccgcag	cgagttgctc	ttgccggcg	9780
tcaacacggg	ataataccgc	gccacatagc	agaactttaa	aagtgctcat	cattggaaaa	9840
gacctgcagg	gggggggggg	cgctgaggtc	tgccctcgta	agaaggtgtt	gctgactcat	9900
accaggcctg	aatcgcccc	tcaccagcc	agaaagtgag	ggagccacgg	ttgatgagag	9960
ctttgttgta	ggtggaccag	ttggtgattt	tgaacttttg	ctttgccacg	gaacggtctg	10020
cgttgtcggg	aagatgcgtg	atctgatcct	tcaactcagc	aaaagttcga	tttattcaac	10080
aaagccgccc	tccgctcaag	tcagcgtaat	gctctgccag	tgttacaacc	aattaacca	10140
ttctgattag	aaaaactcat	cgagcatcaa	atgaaactgc	aatttatca	tatcaggatt	10200
atcaatacca	tatthttgaa	aaagccgtht	ctgtaatgaa	ggagaaaact	caccgaggca	10260
gttccatagg	atggcaagat	cctggtatcg	gtctcgatt	ccgactcgtc	caacatcaat	10320
acaacctatt	aatttcccct	cgtaaaaaat	aaggttatca	agtgagaaat	caccatgagt	10380
gacgactgaa	tccggtgaga	atggcaaaaag	cttatgcatt	tctttccaga	cttgttcaac	10440
agggcagcca	ttacgctcgt	catcaaaatc	actcgcatac	accaaacctg	tattcattcg	10500
tgattgcgcc	tgagcgagac	gaaataccgc	atcgctgtta	aaaggacaat	tacaaacagg	10560
aatcgaatgc	aaccggcgca	ggaacactgc	cagcgcatac	acaatatttt	cacctgaatc	10620
aggatattct	tctaatacct	ggaatgctgt	tttcccgggg	atcgcagttg	tgagtaacca	10680
tgcatcatca	ggagtacgga	taaaatgctt	gatggtcggg	agaggcataa	attccgtcag	10740
ccagthttagt	ctgaccatct	catctgtaac	atcattggca	acgctacctt	tgccatgtht	10800
cagaaacaac	tctggcgcac	cgggcttccc	atacaatcga	tagattgtcg	cacctgattg	10860
cccgcacatta	tcgcgagccc	atthtatccc	atataaatca	gcatccatgt	tggaatttaa	10920
tcgcggcctc	gagcaagacg	tttcccgttg	aatatggctc	ataacacccc	ttgtattact	10980
gthttatgtaa	gcagacagth	ttattgttca	tgatgatata	thtttatctt	gtgcaatgta	11040
acatcagaga	thttgagaca	caacgtggct	thccccccc	cccctgcagg	tcaattcggg	11100
cgatatggct	attacgaaga	aggctcgtgc	gcggagtccc	gtgaactthc	ccacgcaaca	11160
agtgaaaccgc	accgggthtg	ccggaggcca	thtcgttaaa	atgcgcagcc	atggctgctt	11220
cgthccagcat	ggcgtaatac	tgatcctcgt	cttccggctgg	cggtatattg	ccgatgggct	11280
tcaaaaagccg	ccgtgthtga	accagthtat	ccattccaag	gtagcgaact	cgaccgthtc	11340
gaagctcctc	catggtccac	gccgatgaat	gacctcggcc	thgtaaaagac	cgthgatcgc	11400
thctgcgagg	gcgthtctgt	gctgtcgcgc	acgcttcccga	tagatggctc	gatacctgct	11460

-continued

---

tctgccaacc gctcgggaata gcgaaaggac acgtattgaa caccgcgatc cgagtgatgc 11520  
actaggccgc catgagcggg acgccgatca tgatgagcct cctcagagggc atcgaggaca 11580  
aagcctgcat gtgctgtccg gctcgcgccg catccgacaa tgcgacgggc gaagacgtcg 11640  
atcacgaagg ccacgtagac gaagccctcc caagtggcga cataagtacg gacatgcgca 11700  
aaggctttcc cggtttgtcg ctgatggtgc aagagacgct gaagcgcgat ccgatgcgca 11760  
ggcatctggt cgtcttcocg ggtcgtggcg gtggcctgat caaggteact cgccgaagag 11820  
ctgcatgatt ggctcgaaac cgagcggggg aaattgtcgc gcagttctcc cgtcgcgag 11880  
gcgataaatt acatgctcaa gcgatgggat ggcattacgt cattcctcga tgacggcccg 11940  
atctgcctga cgaacaatgc tgcgaacga acgctcagag gctatgtact cggcaggaag 12000  
tcatggctgt ttgccgcatc ggatcgttgt gctgaacgtg cggcgttcat ggcgacactg 12060  
atcatgagcg ccaagctcaa taacatcgat ccgacggcct ggcttgccga cgtccgcgcc 12120  
gaccttgccg acgctccgat cagcaggctt gagcaacagc tgccgtggaa ctggacatcc 12180  
aagacactga gtgctcaggc ggctgacct cggccttca cggataactt accccattat 12240  
cgcagattgc gatgaagcat cagcgtcatt cagcaatctt gccaaagtat gcaggctcgc 12300  
gagaatcgac gtgcgaaacc ggtcgtgtgc gccaaagatc cgcttgccga gcggtcgaa 12360  
atctatgctg ggacttcaag aggtcgagta gaggaagaac cggaaagggt gcaccggaaa 12420  
atatgcgctc ctttgagag cgcctcatgg acgtgaacaa atcgcgccga ccaaggatgc 12480  
cacggataca aaagctcgcg aagctcggtc ccgtgggtgt tctgtcgtct cgttgtataa 12540  
cgaaatccat tcccattccg cgtcaagat ggcttcccct cggcagttca tcagggttaa 12600  
atcaatctag ccgacttgtc cgtgaaatg ggctgcactc caacagaaac aatcaaaaa 12660  
acatacacag cgacttattc acacgagctc aaattacaac ggtatatatc ctgccagtca 12720  
gcatcatcac accaaaagt aggccgaat agtttgaat tagaaagctc gcaattgagg 12780  
tctacaggcc aaattcgtc ttaccgtagc aatattactc accggtgcga tgcccccat 12840  
cgtagggtgaa ggtggaaatt aatgatccat cttgagacca caggcccaca acagctacca 12900  
gtttcctcaa gggccaacca aaaacgtaag cgttacgta catggctgat aagaaaaggc 12960  
aattttaga tgtaacatc caacgtcgtc ttcaggatc gatccaatac gcaaaccgcc 13020  
tctccccgcg cgttgccga ttcattaatg cagctggcac gacaggttc ccgactggaa 13080  
agcgggcagt gagcgcaac caattaatgt gagttagctc actcattagg caccocaggc 13140  
tttacctt atgcttcgg ctcgtatgt gtgtggaatt gtgagcggat aacaatttca 13200  
cacaggaaac agctatgacc atgattacgc caagcttga tgctgcagg tcgactctag 13260  
aggatctggc gcgcccacat ttattatcat ctatttcctg acattttaa ccatccacct 13320  
atgtcaaaaa cttatagaaa atgtcaactt ccaaacaaaa cataattgaa cttcgcaaat 13380  
aaattcttaa taattataa aaatgttact taattattc ttcaacecca tttccgcgc 13440  
gtagcgcgga caaagactc agttaaata agaaagttcc gattctcacc gtataaaacg 13500  
gtgactttgg cgggctttca tgtgtaacaa attggtttaa caaacactg cctagtcggt 13560  
tagttagtaa tcagcgcag gaactccgat tggagcgtga ctttcacgtg ccggaggccc 13620  
accaccacag cggcgcttac gctctaagaa tctcgcacc ggtttcttc atctcccc 13680  
cgccaagtgt ctcctcgtt ccgcaattc catcatgta cagggacct aaaaatggcg 13740  
tatttctca gccccgggta taaatacaca catgatcctg tgggtgggtc ttccacaagt 13800

-continued

---

tacatctcct	tctggTTTT	gtattgcaag	tgTTTgtatt	TTTTgcctcc	gagagaaaat	13860
cgcgccgca	tggagagatc	tcaacggcag	tctcctccgc	caccgtcgcc	gtcctcctcc	13920
tcgTcctccg	tctccgCGGA	caccgtcctc	gtcctcctccg	gaaagaggcg	gagggcgCGG	13980
acggccaagg	ccggcgCGGA	gcctaataag	aggatccgca	aggaccccgC	cgccgCGGcc	14040
gCGGGgaaga	ggagctccgt	ctacagggga	gtcaccaggc	acaggtggac	gggCaggttc	14100
gagcgcatc	tctgggacaa	gcaactgctc	gcccgctcc	acaacaagaa	gaaaggcagG	14160
caagtctaoc	tggggcgGta	tgacagcgag	gaggcagctg	ctcgtgccta	tgacctcgca	14220
gctctcaagt	actggggTcc	tgagactctg	ctcaacttcc	ctgtggagga	ttactccagc	14280
gagatgcccg	agatggaggc	cgtgtcccgg	gaggagtacc	tggcctccct	cgcccgCagG	14340
agcagcggct	tctccagggg	cgtctccaag	tacagaggcg	tcgccaggca	tcaccacaac	14400
gggaggtggg	aggcacggat	tgggcgagtc	tttgggaaca	agtacctcta	cttgggaaca	14460
tttgacactc	aagaagaggc	agccaaggcc	tatgacctg	cgGCCattga	ataccgtggc	14520
gtcaatgctg	taaccaactt	cgacatcagc	tgctaactgg	accacccgct	gttcttgGca	14580
cagctccaac	aggagccaca	ggtggtgCcg	gcaactcaacc	aagaacctca	acctgatcag	14640
agcgaaccg	gaactacaga	gcaagagccg	gagTcaagcg	aagccaagac	accggatggc	14700
agtgcagaac	ccgatgagaa	cgCGgtgCct	gacgacccg	cggagcccct	caccacagtc	14760
gacgacagca	tcgaagaggg	cttgtggagc	ccttgcatgg	attacgagct	agacaccatg	14820
tcgagaccaa	actttggcag	ctcaatcaat	ctgagcgagT	ggttcgctga	cgCagacttc	14880
gactgcaaca	tcggatgCct	gttcgatggg	tgTtctgCgg	ctgacgaagg	aagcaaggat	14940
ggtgtaggtc	tggcagatTt	cagtctgTtt	gaggcaggtg	atgtccagct	gaaggatgTt	15000
ctttcggata	tggaaagggg	gatacaacct	ccagcgatga	tcagtgtgtg	caacgCGGcc	15060
gcaagtatga	actaaaatgc	atgtaggtgt	aagagctcat	ggagagcatg	gaatattgta	15120
tccgaccatg	taacagtata	ataactgagc	tccatctcac	ttcttctatg	aataaacaaa	15180
ggatgttatg	atatattaac	actctatcta	tgcaacctat	tgTtctatga	taaaTttcct	15240
cttattatta	taaTcatctc	gaatcgtgac	ggcttatgga	atgcttcaaa	tagtcaaaaa	15300
acaaatgtgt	actataagac	tttctaaca	attctaacct	tagcattgtg	aacgagacat	15360
aagTgttaag	aagacataac	aattataatg	gaagaagTtt	gtctccattt	atatattata	15420
tattaccocac	ttatgtatta	tattaggatg	ttaaggagac	ataacaatta	taaagagaga	15480
agTttgtatc	catttatata	ttatatacta	cccatttata	tattatactt	atccacttat	15540
ttaatgtctt	tataaggTtt	gatccatgat	atTtctaata	TTTTagttga	tatgtatatg	15600
aaagggtact	atTtgaactc	tcttactctg	tataaaggTt	ggatcatcct	taaagTgggt	15660
ctatttaatt	ttattgcttc	ttacagataa	aaaaaaaaatt	atgagTtggT	ttgataaaat	15720
attgaaggat	ttaaataaat	aataaataac	atataatata	tgTatataaa	tttattataa	15780
tataacattt	atctataaaa	aagTaaatat	tgTcataaat	ctatacaatc	gTttagcctt	15840
gctggacgaa	tctcaattat	ttaaacgaga	gTaaacatat	ttgactTttt	ggttatTtaa	15900
caaattatta	tttaacacta	tatgaaatTt	TTTTTTTTat	cagcaaaGaa	taaaattaaa	15960
ttaagaagga	caatggTgtc	ccaatcctta	tacaaccaac	ttccacaaga	aagTcaagtc	16020
agagacaaca	aaaaacaag	caaaggaaat	TTTTtaattt	gagTtgtctt	gTttgctgca	16080
taatttatgc	agTaaaacac	tacacataac	cctTttagca	gtagagcaat	ggtTgaccgt	16140
gtccttagct	tctTttatTt	tattTttTta	tcagcaaaGaa	ataaataaaa	taaaatgaga	16200

-continued

---

cacttcaggg atgtttcaac aagcttggat cctcgaagag aagggttaat aacacacttt 16260  
ttaaaccattt ttaacacaaa ttttagttat ttaaaaaattt attaaaaaat ttaaaataag 16320  
aagaggaact ctttaaataa atctaactta caaaatttat gatttttaat aagttttcac 16380  
caataaaaaa tgtcataaaa atatgttaaa aagtatatta tcaatattct ctttatgata 16440  
aataaaaaga aaaaaaaaaat aaaagttaag tgaaaatgag attgaagtga ctttaggtgt 16500  
gtataaatat atcaaccccg ccaacaattt atttaacca aatataatga agtatattat 16560  
tccatagcct ttattttattt atatatttat tatataaaag ctttatttgt tctaggttgt 16620  
tcatgaaata tttttttggt tttatctcgg ttgtaagaaa atcatgtgct ttgtgtcgcc 16680  
actcactatt gcagcttttt catgcattgg tcagattgac gggtgattgt atttttgttt 16740  
tttatggttt tgtgttatga cttaagtctt catctcttta tctcttcate aggtttgatg 16800  
gttacctaata atggctcatg ggtacatgca tggttaaatt aggtggccaa ctttgttgtg 16860  
aacgatagaa ttttttttat attaagtaaa ctatttttat attatgaaat aataataaaa 16920  
aaaatatttt atcattatta acaaaatcat attagttaat ttgttaactc tataataaaa 16980  
gaaatactgt aacattcaca ttacatggtg acatctttcc accctttcat ttgttttttg 17040  
tttgatgact tttttcttg tttaaattta tttcccttct tttaaatttg gaatacatta 17100  
tcatcatata taaactaaaa tactaaaaac aggattacac aaatgataaa taataacaca 17160  
aatatttata aatctagctg caatatattt aaactagcta tatcgatatt gtaaaataaa 17220  
actagctgca ttgatactga taaaaaaata tcatgtgctt tctggactga tgatgcagta 17280  
tacttttgac attgccttta ttttattttt cagaaaagct ttcttagttc tgggttcttc 17340  
attatttgtt tcccatctcc attgtgaatt gaatcatttg cttcgtgtca caaatacaat 17400  
ttagntaggt acatgcattg gtcagattca cggtttatta tgatcatgact taagttcatg 17460  
gtagtacatt acctgccacg catgcattat attggttaga tttgatagge aaatttgggt 17520  
gtcaacaata taaatataaa taatgttttt atattacgaa ataacagtga tcaaaacaaa 17580  
cagttttatc tttattaaca agattttgtt tttgtttgat gacgtttttt aatgtttacg 17640  
ctttccctct tcttttgaat ttagaacact ttatcatcat aaaatcaaat actaaaaaaaa 17700  
ttacatattt cataaataat aacacaaata tttttaaaaa atctgaaata ataataaaca 17760  
atattacata ttatcacgaa aattcattaa taaaaatatt atataataaa aatgtaatag 17820  
tagttatatg taggaaaaaa gtactgcacg cataatatat acaaaaagat taaaatgaac 17880  
tattataaat aataacacta aattaatggt gaatcatatc aaaataatga aaaagtaaat 17940  
aaaatttght attaaactct atatgtatta cacacacaaa taataaataa tagtaaaaaa 18000  
aattatgata aatatttacc atctcataag atatttataa taatgataaa aatatagatt 18060  
atthtttatg caactagcta gccaaaaaga gaacacgggt atatataaaa agagtacctt 18120  
taaattctac tgtacttctt ttattcttga cgtttttata tcaagtggac atacgtgaag 18180  
atthtaatta tcagtctaaa tatttcatta gcaacttaata cttttctggt ttattcttat 18240  
cctataagta gtcccgatc tcccaacatt gcttattcac acaactaact aagaaagtct 18300  
tccatagccc cccaagcggc ccatggctc ctcgaggac gtcacaaagg agttcatgag 18360  
cttcaaggtg cgcattggagg gctccgtgaa cggccacgag ttcgagatcg agggcgaggg 18420  
cgagggccgc ccctacgagg gccccagac cgccaagctg aaggtgacca agggcgggcc 18480  
cctgcccctc gcctgggaca tctgtcccc ccagttccag tacggctcca aggtgtacgt 18540

-continued

---

```

gaagcacccc gccgacatcc ccgactacaa gaagctgtcc ttccccgagg gcttcaagtg 18600
ggagcgcgtg atgaacttcg aggacggcgg cgtggtgacc gtgaccagg actcctcctc 18660
gcaggacggc tccttcatct acaagtgtaa gtccatcggc gtgaacttcc cctccgacgg 18720
ccccgtaatg cagaagaaga ctatgggtg ggaggcctcc accgagcggc tgtacccccg 18780
cgacggcgtg ctgaagggcg agatccacaa ggcctgaag ctgaaggacg gcggccacta 18840
cctggtggag ttcaagtcca tctacatggc caagaagccc gtgcagctgc ccggtacta 18900
ctacgtggac tccaagctgg acatcacctc ccacaacgag gactacacca tcgtggagca 18960
gtacgagcgc gccgagggcc gccaccacct gttcctgtag cggccggcgg cgacacaagt 19020
gtgagagtac taaataaatg ctttggttgt acgaaatcat taccataaat aaaataatca 19080
aagcttatat atgccttcgg ctaaggccga atgcaaagaa attggttctt tctcgttatc 19140
ttttgccact tttactagta cgtattaatt actactaat catctttgtt tacggetcat 19200
tatatccgtc gacggg                                     19215

```

```

<210> SEQ ID NO 67
<211> LENGTH: 9565
<212> TYPE: DNA
<213> ORGANISM: Artificial sequence
<220> FEATURE:
<223> OTHER INFORMATION: Vector KS428
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: (4330)..(4330)
<223> OTHER INFORMATION: n is a, c, g, or t
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: (7746)..(7746)
<223> OTHER INFORMATION: n is a, c, g, or t

```

```

<400> SEQUENCE: 67

```

```

ctagagggcc caattcggcc tatagtgagt cgtattacaa ttcaactggcc gtcgttttac 60
aacgtcgtga ctgggaaaac cctggcggtta cccaacttaa tcgaccttga gcacatcccc 120
ctttcgccag ctggcgtaat agcgaagagg cccgcaccga tcgaccttcc caacagttgc 180
gcagcctata cgtacggcag ttaaggttt acacctataa aagagagagc cgttatcgtc 240
tgtttggtga tgtacagagt gatattattg acacgccggg gcgacggatg gtgatcccc 300
tggccagtgc acgtctcgtg tcagataaag tctcccgta actttaccgg gtggtgcata 360
tcggggatga aagctggcgc atgatgacca ccgatatggc cagtgtgccg gtctccgtta 420
tcggggaaga agtggctgat ctacagccacc gcgaaaatga catcaaaaac gccattaacc 480
tgatgttctg gggaaataaa atgtcaggca tgagattatc aaaaaggatc ttcacctaga 540
tccttttcac gtagaaaagc agtccgcaga aacggtgctg accccggatg aatgtcagct 600
actgggctat ctggacaagg gaaaacgcaa gcgcaaagag aaagcaggta gcttgacagt 660
ggcttacatg gcgatagcta gactggggcg ttttatggac agcaagcgaa ccggaattgc 720
cagctggggc gccctctggt aaggttggga agccctgcaa agtaaaactgg atggctttct 780
tgccgccaaag gatctgatgg cgcaggggat caagctctga tcaagagaca ggatgaggat 840
cgtttcgcac gattgaacaa gatggattgc acgcaggttc tccggccgct tgggtggaga 900
ggctattcgg ctatgactgg gcacaacaga caatcggctg ctctgatgcc gccgtgttcc 960
ggctgtcagc gcagggggcg ccggttcttt ttgtcaagac cgacctgtcc ggtgccctga 1020
atgaactgca agacgaggca gcgcggtat cgtggctggc cacgacgggc gttccttgcg 1080
cagctgtgct cgacgttctc actgaagcgg gaagggactg gctgctattg ggcgaagtgc 1140

```

-continued

---

cggggcagga tctcctgtca tctcaccttg ctctcgccga gaaagtatcc atcatggctg	1200
atgcaatgcg gcggctgcat acgcttgatc cggtacctg cccattcgac caccaagcga	1260
aacatcgcat cgagcgagca cgtactcggg tggagccgg tcttgctgat caggatgatc	1320
tggacgaaga gcatcagggg ctgcgcccag ccgaactggt cggcaggctc aaggcgagca	1380
tgcccagcgg cgaggatctc gtcgtgacc atggcgatgc ctgcttgccg aatatcatgg	1440
tggaaaatgg ccgcttttct ggattcatcg actgtggccg gctgggtgtg gcggaccgct	1500
atcaggacat agcgttggtt acccgtgata ttgctgaaga gcttgccggc gaatgggctg	1560
accgcttctt cgtgctttac ggtatcgccg ctcccgatc gcagcgcac gccttctatc	1620
gccttcttga cgagtcttct tgaattatta acgcttaca tttctgatg cggatatttc	1680
tccttacgca tctgtgcggt atttcacacc gcatcaggtg gcacttttct gggaaatgtg	1740
cgcggaacc ctatttgttt atttttctaa atacattcaa atatgatcc gctcatgaga	1800
caataaacct gataaatgct tcaataatag cacgtgagga gggccaccat ggccaagttg	1860
accagtgcgg ttccgggtgt caccgcgcgc gacgtcgccg gagcggctga gttctggacc	1920
gaccggctcg ggttctcccg ggactctgtg gaggacgact tcgcccgtgt ggtccgggac	1980
gacgtgacc tgttcatcag cgcggtccag gaccagggtg tgccggacaa caccctggcc	2040
tgggtgtggg tgccggcctt ggacgagctg tacgcccagt ggtcggaggt cgtgtccacg	2100
aaactccggg acgctccggg gccggccatg accgagatcg gcgagcagcc gtggggcggg	2160
gagttcgcgc tgccgcagcc gccggcaac tgcgtgcact tcgtggccga ggagcaggac	2220
tgacacgtgc taaaacttca tttttaattt aaaaggatct aggtgaagat cctttttgat	2280
aatctcatga ccaaaatccc ttaacgtgag ttttcgttcc actgagcgtc agaccocgta	2340
gaaaagatca aaggatcttc ttgagatcct ttttttctgc gcgtaatctg ctgcttgcaa	2400
acaaaaaac caccgctacc agcgggtggt tgtttgcggg atcaagagct accaactctt	2460
tttccgaagg taactggctt cagcagagcg cagataccaa atactgttct tctagtgtag	2520
ccgtagttag gccaccactt caagaactct gtagcaccgc ctacatacct cgctctgcta	2580
atcctgttac cagtggctgc tgccagtggc gataagctgt gtcttaccgg gttggactca	2640
agacgatagt taccggataa ggcgcagcgg tcgggctgaa cggggggttc gtgcacacag	2700
cccagcttgg agcgaacgac ctacaccgaa ctgagatacc tacagcgtga gctatgagaa	2760
agcgcaccgc ttcccgaagg gagaaaggcg gacaggtatc cggtaagcgg cagggtcggg	2820
acaggagagc gcacgagggg gcttccaggg ggaacgcct ggtatcttta tagtcctgtc	2880
gggtttcgcc acctctgact tgagcgtcga tttttgtgat gctcgtcagg ggggcgggac	2940
ctatggaaaa acgcccagaa cgcggccttt ttacgggtcc tggccttttg ctggcctttt	3000
gctcacatgt tctttcctgc gttatcccct gattctgtgg ataaccgtat tacgccttt	3060
gagtgagctg ataccgctcg ccgcagcccga acgaccgagc gcagcagctc agtgagcgag	3120
gaagcgggag agcgcaccaat acgcaaacgg cctctcccgg cgcgttgggc gattcattaa	3180
tgcagctggc acgacaggtt tcccagctgg aaagcgggca gtgagcgcga cgcaattaat	3240
gtgagttagc tcaactatta ggcaccccag gctttacact ttatgcttcc ggctcgtatg	3300
ttgtgtggaa ttgtgagcgg ataacaattt cacacaggaa acagctatga ccatgattac	3360
gccaagctat ttaggtagc cgttagaata ctcaagctat gcatcaagct tggtagcgag	3420
ctcgatcca ctagtaacgg ccgcccgtgt gctggaatc aggggcgcgc cctatagatg	3480

-continued

---

ggatgaagct gctctcgaca aatctgataa aactaaagaa ggtagtaat caatttttac	3540
aaaatcatag attatTTTTt tcattgaatt atTTTTatgc tataccaaga attgtatTTT	3600
agtatttTgt ttaactacat ataatagaat taactacata taaattaact aaacttaaaa	3660
taaaaataga tttgttTcoct gaaattattt taagaatata tatgtatata tctaaaaTct	3720
tagacttaga tagattttTtc tatctatctta ttttggttac ttaaaataaa taaattTgta	3780
taaataattg tatagttatc aaaaattaaa actaattttt ttaaagtTgt tgatatataa	3840
aatactaaag atttaacgat taagtattta tttaagtata gaatttTgtt ttttttttaa	3900
gttagttat gaagtTgtta attatattaa aacaaaacaa tatttcgaaa ttttattatc	3960
atattcgaat atattttttt tagtgatgat gtatgaatta ttatcataat ttgaaagtTt	4020
actaaaaat atatcaacat gaattgtaat atatgagtta ttacctaac caaaattata	4080
aattaacatt aaatataatt atatatgtca tatttagcca tacaatTgt catcaatatt	4140
aatagtcag tcaatattac ataatgocaa tattatgcta cttaaaccoc aaatccoccta	4200
actcccgTta agtagccaaa ttcataaata tacttattcg acaaaataaa aaactttaaa	4260
atatttacta atccgacat gcacaagcat ccattcccta ttccattgcc acgggataac	4320
aatgcaaccn actcctcaaa aaaagaaaaa ttcaagctct tttgcaaaaa aaaaataaat	4380
aattttaaca cctaaaattt tttgttTcca aacttctaca gggaacacac ataaaagaaa	4440
aagaggacgt ccactcggat cacgcaacaa accaaaaggt gtgtcatgac tcctaagata	4500
taatatTtcc ttattcaaaa tcataccatt ttaaattatg aatgtatttc gtagtccacc	4560
agatatgTaa tccaccagcg ttcaaoccaa agttttatga ttgtaagtTt aagtgaatta	4620
taataatata ttcttcacgg tatcttttca taaactaattg agttatcaaa cttgatcgca	4680
catgtggctt tgataggtgt gacttttatg gtatacaatt ctttcaacct aaaaacatta	4740
ttgttccTca atatcttaca ttatgctTga ctgcaacaaa atattttctc atctgttttc	4800
ttcctttaaa ccaatttatt atcatctatt tcttgacatt ttaatccatc cacctatgtc	4860
aaaaacttat agaaaatgtc aacttccaaa caaaacataa tggaacttcg caaataaatt	4920
cttaataata ttaaaaaatg ttacttaatt atttcttcaa ccccatTTTc cgcgcgtagc	4980
gcgacaaaag actctagtta aatatagaag ttcccgattc tcatcgtata aaacggtgac	5040
tttgccgggc tttcatgtgt aacaaattgg tttaacaaac cactgcctag tctgttagtg	5100
tagaatcagc gcatggaact ccgattggag cgtgactTtc acgtgccgga ggcccaccac	5160
cacagccggc gttacgctct aagaatctcg cccacggTtt tcttcatctg ccccccgcca	5220
agtgctTtcc tctgtcgcca cttctcacca agttacagga accctaaaaa tggcctttct	5280
tcagccccgg ctataatata cacatgatcc tatagtgggt tcttccacaa gttacatctc	5340
cttctggatt gtacatttca agtgttTgtg ttttttctgc ctctgagaga aaatcgcggc	5400
cgcaagtatg aactaaaatg catgtaggtg taagagctca tggagagcat ggaatattgt	5460
atccgacat gtaacagtat aataactgag ctccatctca cttcttctat gaataaacia	5520
aggatgttat gatataTtaa cactctatct atgcacctta ttgttctatg ataaattTcc	5580
tcttattatt ataaatcatc tgaatcgtga cggcttatgg aatgctTcaa atagtacaaa	5640
aacaaatgtg tactataaga ctttctaaac aattctaacc ttagcattgt gaacgagaca	5700
taagtgtTaa gaagacataa caattataat ggaagaagtt tgtctccatt tatatattat	5760
atattaccCa cttatgtatt atattaggat gtttaaggaga cataacaatt ataaagagag	5820
aagttTgtat ccatttatat attatatact acccatttat atattatact tatccactta	5880

-continued

---

ttaaagtct	ttataaggtt	tgatecatga	tatttcta	atattagttg	atatgtatat	5940
gaaaggtac	tatttgaact	ctcttactct	gtataaaggt	tggatcatcc	ttaaagtggg	6000
tctattta	tttattgctt	cttacagata	aaaaaaaa	tatgagttgg	tttgataaaa	6060
tattgaagga	tttaaaataa	taataaataa	catataatat	atgtatataa	atattattata	6120
atataacatt	tatctataaa	aaagtaata	ttgtcataaa	tctatacaat	cgtttagcct	6180
tgctggacga	atctcaatta	tttaaagcag	agtaaacata	tttgactttt	tggttattta	6240
acaaattatt	atttaacact	atatgaaatt	ttttttttta	tcagcaaaga	ataaaattaa	6300
attaagaagg	acaatggtgt	ccaatcctt	atacaaccaa	cttccacaag	aaagtcaagt	6360
cagagacaac	aaaaaaaa	gcaaaggaaa	ttttttaatt	tgagttgtct	tgtttctgct	6420
ataatttatg	cagtaaaaca	ctacacataa	cccttttagc	agtagagcaa	tggttgaccg	6480
tgctcttagc	ttcttttatt	ttattttttt	atcagcaaag	aataaataaa	ataaaatgag	6540
acacttcagg	gatgtttcaa	caagcttgg	tctcgaaga	gaagggtaa	taacacactt	6600
ttttaacatt	tttaacacaa	attttagtta	tttaaaatt	tattaaaaaa	tttaaaataa	6660
gaagaggaac	tctttaaata	aatctaactt	acaaaattta	tgatttttaa	taagttttca	6720
ccaataaaaa	atgtcataaa	aatatgttaa	aaagtatatt	atcaatattc	tctttatgat	6780
aaataaaaag	aaaaaaaa	taaaagttaa	gtgaaaatga	gattgaagtg	actttaggtg	6840
tgataaata	tatcaacccc	gccaacaatt	tatttaatcc	aaatatattg	aagtatatta	6900
ttccatagcc	tttatttatt	tatatattta	ttatataaaa	gctttatttg	ttctaggttg	6960
ttcatgaaat	atttttttgg	ttttatctcc	gttgaagaa	aatcatgtgc	tttgtgtcgc	7020
cactcactat	tcagcctttt	tcatgcattg	gtcagattga	cggttgattg	tatttttgtt	7080
ttttatggtt	ttgtgttatg	acttaagtct	tcctctcttt	atctcttcat	caggtttgat	7140
ggttacctaa	tatggtccat	gggtacatgc	atggttaaat	taggtggcca	actttgttgt	7200
gaacgataga	atttttttta	tattaagtaa	actattttta	tattatgaaa	taataataaa	7260
aaaaatattt	tatcattatt	aacaaaatca	tattagttaa	tttgtttaact	ctataataaa	7320
agaaatactg	taacattcac	attacatggt	aacatctttc	caccctttca	tttgtttttt	7380
gtttgatgac	ttttttctt	gtttaaattt	atttcccttc	ttttaattt	ggaatacatt	7440
atcatcatat	ataaactaaa	atactaaaaa	caggattaca	caaatgataa	ataataacac	7500
aaatatttat	aaatctagct	gcaatatatt	taaactagct	atatcgatat	tgtaaaataa	7560
aactagctgc	attgatactg	ataaaaaaat	atcatgtgct	ttctggactg	atgatgcagt	7620
atacttttga	cattgccttt	attttatttt	tcagaaaagc	tttcttagtt	ctgggttctt	7680
cattattttg	ttcccatctc	cattgtgaat	tgaatcattt	gcttcgtgtc	acaaatacaa	7740
tttagntagg	tacatgcatt	ggtcagattc	acggtttatt	atgtcatgac	ttaagttcat	7800
ggtagtacat	tacctgccac	gcctgcatta	tattggttag	atttgatagg	caaatgtggt	7860
tgtaacaat	ataaatataa	ataatgtttt	tatattacga	aataacagtg	atcaaaaaca	7920
acagttttat	ctttattaac	aagattttgt	ttttgtttga	tgacgttttt	taatgtttac	7980
gctttcccc	ttcttttgaa	tttagaacac	tttatcatca	taaaatcaaa	tactaaaaaa	8040
attacatatt	tcataaataa	taacacaaat	attttataaa	aatctgaaat	aataatgaac	8100
aatattacat	attatcacga	aaattcatta	ataaaaaat	tatataaata	aaatgtaata	8160
gtagttatat	gtaggaaaaa	agtactgcac	gcataatata	tacaaaaaga	ttaaaatgaa	8220

-continued

---

```

ctattataaa taataacact aaattaatgg tgaatcatat caaataatg aaaaagtaaa 8280
taaaatttgt aattaacttc tatatgtatt acacacacaa ataataaata atagtaaaaa 8340
aaattatgat aaatatttac catctcataa gatattttaa ataatgataa aaatatagat 8400
tattttttat gcaactagct agccaaaaag agaacacggg tatatataaa aagagtacct 8460
ttaaattcta ctgtacttcc tttattcctg acgtttttat atcaagtgga catacgtgaa 8520
gattttaatt atcagtctaa atatttcatt agcacttaat acttttctgt tttattccta 8580
tcctataagt agtcccgatt ctcccaacat tgcttattca cacaactaac taagaaagtc 8640
ttccatagcc ccccaagcgg cccatggcct cctccgagga cgtcatacaag gagttcatgc 8700
gcttcaaggt gcgcatggag ggtcctgtga acggccacga gttecgagatc gagggcgagg 8760
gcgagggcgg cccctacgag ggcaccocaga ccgccaagct gaaggtgacc aagggcggcc 8820
ccctgccctt cgctgggac atcctgtccc cccagttcca gtaaggctcc aagggtgtacg 8880
tgaagcacc cgcgcacatc ccgcactaca agaagctgtc cttccccgag ggcttcaagt 8940
gggagcgcgt gatgaacttc gaggacggcg gcgtgggtgac cgtgaccag gactcctccc 9000
tgcaggacgg ctccttcac tacaaggtga agttcatcgg cgtgaacttc ccctccgacg 9060
gccccgtaat gcagaagaag actatgggct gggaggcctc caccgagcgc ctgtaccccc 9120
gcgacggcgt gctgaagggc gagatccaca aggccctgaa gctgaaggac ggcggccact 9180
acctgggtgga gttcaagtcc atctacatgg ccaagaagcc cgtgcagctg cccggctact 9240
actacgtgga ctccaagctg gacatcacct cccacaacga ggactacacc atcgtggagc 9300
agtacgagcg cgccgagggc cgccaccacc tgttcctgta gcggccggcc gcgacacaag 9360
tgtgagagta ctaataaat gctttggtg tacgaaatca ttacactaaa taaaataatc 9420
aaagcttata tatgccttcc gctaaggcgg aatgcaaaaga aattggttct ttctcgttat 9480
cttttgccac ttttactagt acgtattaat tactacttaa tcatctttgt ttacggctca 9540
ttatatccgt cgacggcgcg ccgct 9565

```

```

<210> SEQ ID NO 68
<211> LENGTH: 10947
<212> TYPE: DNA
<213> ORGANISM: Artificial sequence
<220> FEATURE:
<223> OTHER INFORMATION: Vector KS429
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: (2349)..(2349)
<223> OTHER INFORMATION: n is a, c, g, or t
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: (8498)..(8498)
<223> OTHER INFORMATION: n is a, c, g, or t

```

```

<400> SEQUENCE: 68

```

```

ggccgcaagt atgaactaaa atgcatgtag gtgtaagagc tcatggagag catggaatat 60
tgtatccgac catgtaacag tataataact gagctccatc tcacttcttc tatgaataaa 120
caaaggatgt tatgatatat taacactcta tctatgcacc ttattgttct atgataaatt 180
tcctcttatt attataaatc atctgaatcg tgacggccta tggaatgctt caaatagtac 240
aaaaacaaat gtgtactata agactttcta aacaattcta accttagcat tgtgaacgag 300
acataagtgt taagaagaca taacaattat aatggaagaa gtttgtctcc atttatatat 360
tatatattac ccacttatgt attatattag gatgtaagg agacataaca attataaaga 420
gagaagtttg tatccattta tatattatat actaccatt tatatattat acttatccac 480

```

-continued

---

ttatattaag tctttataag gtttgatcca tgatatttct aatatttttag ttgatatgta	540
tatgaaaggg tactatttga actctcttac tctgtataaa gggtggatca tccttaaagt	600
gggtctatatt aattttattg cttcttacag ataaaaaaaa aattatgagt tggtttgata	660
aaatattgaa ggatttataa taataataaa taacatataa tatatgtata taaatttatt	720
ataatataac atttatctat aaaaaagtaa atattgtcat aaatctatac aatcgtttag	780
ccttgctgga cgaatctcaa ttatttaaac gagagtaaac atatttgact ttttggttat	840
ttaacaaatt attatttaac actatatgaa attttttttt ttatcagcaa agaataaaat	900
taaattaaga aggacaatgg tgtccaatc cttatacaac caactccac aagaaagtca	960
agtcagagac aacaaaaaaaa caagcaagg aaatttttta atttgagttg tcttgtttgc	1020
tgcataatatt atgcagtaaa acactacaca taacctttt agcagtagag caatggttga	1080
ccgtgtgctt agcttctttt attttatttt tttatcagca aagaataaat aaaataaaat	1140
gagacacttc agggatgttt caacaagctt ggatcctcga agagaagggt taataacaca	1200
cttttttaac attttttaaca caaatttttag ttatttaaaa atttattaaa aaatttataa	1260
taagaagagg aactctttaa ataaatctaa cttacaaaat ttatgatatt taataagttt	1320
tcaccaataa aaaatgtcat aaaaatattg taaaaagtat attatcaata ttctctttat	1380
gataaataaa aagaaaaaaaa aaataaaaagt taagtgaaaa tgagattgaa gtgactttag	1440
gtgtgtataa atatatcaac cccgccaaca atttatttaa tccaaatata ttgaagtata	1500
ttattccata gcctttatatt atttatata ttattatata aaagctttat ttgttctag	1560
ttgttcatga aatatttttt tggttttatc tccgttgtaa gaaaatcatg tgccttgggt	1620
cgccactcac tattgcagct tttcatgca ttggtcagat tgacggttga ttgtatttt	1680
gttttttatg gttttgtgtt atgacttaag tcttcatctc tttatctctt catcaggttt	1740
gatggttacc taatatggtc catgggtaca tgcattggtta aattaggtgg ccaactttgt	1800
tgtgaacgat agaatttttt ttatattaag taaactattt ttatattatg aaataataat	1860
aaaaaaaaata ttttatcatt attaacaaaa tcatattagt taatttgta actctataat	1920
aaaagaaata ctgtaacatt cacattacat ggtaacatct tccaccctt tcatttgttt	1980
tttgtttgat gacttttttt cttgtttaa tttatttccc tctttttaa tttggaatac	2040
attatcatca tataaaact aaaatactaa aaacaggatt acacaaatga taaataataa	2100
cacaaatatt tataaatcta gctgcaatat atttaaacta gctatatcga tattgtaaaa	2160
taaaactagc tgcattgata ctgataaaaa aatatcatgt gctttctgga ctgatgatgc	2220
agtatacttt tgacattgcc tttattttat ttttcagaaa agctttctta gttctgggtt	2280
cttcattatt tgtttcccat ctccattgtg aattgaatca tttgcttctg gtcacaaata	2340
caatttagnt aggtacatgc attggtcaga ttcacggttt attatgtcat gacttaagtt	2400
catggtagta cattacctgc cacgcacgca ttatattggt tagatttgat aggcaaat	2460
ggttgtcaac aatataaata taaataatgt ttttatatta cgaaataaca gtgatcaaaa	2520
caaacagttt tatctttatt aacaagattt tgtttttgtt tgatgacgtt ttttaagtt	2580
tacgctttcc cccttctttt gaatttagaa cactttatca tcataaaatc aaatactaaa	2640
aaaattacat atttcataaa taataacaca aatattttta aaaaatctga aataataatg	2700
aacaatatta catattatca cgaaaattca ttaataaaaa tattatataa ataaaatgta	2760
atagtagtta tatgtaggaa aaaagtactg cacgcataat atatacaaaa agattaaaat	2820

-continued

---

gaactattat	aaataataac	actaaattaa	tggtgaatca	tatcaaaata	atgaaaaagt	2880
aaataaaaatt	tgtaattaac	ttctatatgt	attacacaca	caaataataa	ataatagtaa	2940
aaaaaattat	gataaatatt	taccatctca	taagatattt	aaaataatga	taaaaatata	3000
gattatTTTT	tatgcaacta	gctagccaaa	aagagaacac	gggtatatat	aaaaagagta	3060
cctttaaatt	ctactgtact	tcctttattc	ctgacgtttt	tatatcaagt	ggacatacgt	3120
gaagatttta	attatcagtc	taaatatttc	attagcactt	aatacttttc	tgtttttattc	3180
ctatcctata	agtagtcccg	attctcccaa	cattgtttat	tcacacaact	aactaagaaa	3240
gtcttcata	gcccccaag	cggcccatgg	cctcctccga	ggacgtcatc	aaggagtcca	3300
tgcgcttcaa	ggtgcgcatg	gagggctccg	tgaacggcca	cgagttcgag	atcgagggcg	3360
agggcgaggg	ccgcccctac	gagggcacc	agaccgcaa	gctgaagggtg	accaagggcg	3420
gccccctgcc	cttcgctcgg	gacatcctgt	ccccccagtt	ccagtaaggc	tccaagggtg	3480
acgtgaagca	ccccgcccac	atccccgact	acaagaagct	gtccttcccc	gagggcttca	3540
agtgggagcg	cgtgatgaac	ttcgaggacg	gcggcgtgg	gaccgtgacc	caggactcct	3600
ccctgcagga	cggctccttc	atctacaagg	tgaagttcat	cgcggtgaac	ttcccctccg	3660
acggccccgt	aatgcagaag	aagactatgg	gctggggaggc	ctccaccgag	cgctgtacc	3720
cccgcgacgg	cgtgctgaag	ggcgagatcc	acaaggcct	gaagctgaag	gacggcgcc	3780
actacctggt	ggagtccaag	tccatctaca	tggccaagaa	gcccgtgcag	ctgccccgct	3840
actactacgt	ggactccaag	ctggacatca	cctcccacaa	cgaggactac	accatcgtgg	3900
agcagtacga	gcgcgcccag	ggccgcccacc	acctgttct	gtagcggccg	gccgcgacac	3960
aagtgtgaga	gtactaaata	aatgctttgg	ttgtacgaaa	tcattacact	aaataaaata	4020
atcaaaagctt	atatatgctt	tccgctaagg	cgaatgcaa	agaaattggt	tctttctcgt	4080
tatcttttgc	cacttttact	agtacgtatt	aattactact	taatcatctt	tgtttacggc	4140
tcattatatac	cgtcgaaggc	gcgcgctct	agagggccca	attcgcccta	tagtgagtcg	4200
tattacaatt	cactggccgt	cgttttacia	cgctcgtgact	gggaaaacc	tggcgttacc	4260
caacttaatc	gccttgacgc	acatccccct	ttcgccagct	ggcgtaatag	cgaagaggcc	4320
cgaccgac	gccccttcca	acagttgcgc	agcctatacg	tacggcagtt	taaggtttac	4380
acctataaaa	gagagagccg	ttatcgtctg	ttgtggatg	tacagagtga	tattattgac	4440
acgcccgggc	gacggatggt	gatccccctg	gccagtgcac	gtctgctgtc	agataaagtc	4500
tcccgtgaac	tttaccgggt	ggtgcatatc	ggggatgaaa	gctggcgcac	gatgaccacc	4560
gatatggcca	gtgtgccggt	ctccgttata	ggggaagaag	tggctgatct	cagccaccgc	4620
gaaaatgaca	tcaaaaacgc	cattaacctg	atgttctggg	gaatataaat	gtcaggcatg	4680
agattatcaa	aaaggatcct	cacctagatc	cttttcacgt	agaaagccag	tccgcagaaa	4740
cggctctgac	cccggatgaa	tgtcagctac	tgggctatct	ggacaaggga	aaacgcaagc	4800
gcaaagagaa	agcaggtagc	ttgcagtggg	cttacatggc	gatagctaga	ctgggcggtt	4860
ttatggacag	caagcgaacc	ggaattgcca	gctggggcgc	cctctggtaa	ggttgggaag	4920
ccctgcaaag	taaactggat	ggctttcttg	ccgccaagga	tctgatggcg	caggggatca	4980
agctctgac	aagagacagg	atgaggatcg	tttcgcatga	ttgaacaaga	tggattgcac	5040
gcaggttctc	cggccgcttg	ggtggagagg	ctattcggct	atgactgggc	acaacagaca	5100
atcgctgct	ctgatgccgc	cgtgttccgg	ctgtcagcgc	aggggcgccc	ggttcttttt	5160
gtcaagacgc	acctgtccgg	tgccctgaat	gaactgcaag	acgaggcagc	gcggctatcg	5220

-continued

---

tggtggcca	cgacgggogt	tccttgogca	gctgtgctcg	acgttgccac	tgaagcggga	5280
agggactggc	tgctattggg	cgaagtgcg	gggcaggatc	tcctgtcacc	tcaccttget	5340
cctgocgaga	aagtatccat	catggctgat	gcaatgcggc	ggctgcatac	gcttgatccg	5400
gctacctgcc	cattcgacca	ccaagcgaaa	catcgcatcg	agcgagcacg	tactcggatg	5460
gaagccggtc	ttgtcgatca	ggatgatctg	gacgaagagc	atcaggggct	cgccgagcc	5520
gaactgttcg	ccaggctcaa	ggcgagcatg	cccgaocggc	aggatctcgt	cgtgacccat	5580
ggcgatgcct	gcttgccgaa	tatcatggtg	gaaaatggcc	gcttttctgg	attcatcgac	5640
tgtggccggc	tgggtgtggc	ggaccgctat	caggacatag	cggtggctac	ccgtgatatt	5700
gctgaagagc	ttggcggcga	atgggctgac	cgcttcctcg	tgctttacgg	tatcgccgct	5760
cccgatctgc	agcgcatcgc	cttctatcgc	cttcttgacg	agttctctcg	aattattaac	5820
gcttacaatt	tcctgatgcg	gtatttctc	cttacgcacc	tgtgcgggat	ttcacaccgc	5880
atcaggtggc	acttttcggg	gaaatgtcgc	cggaaacctc	atgtgttat	ttttctaaat	5940
acattcaaat	atgtatccgc	tcatgagaca	ataacctga	taaatgcttc	aataatagca	6000
cgtgaggagg	gccaccatgg	ccaagttgac	cagtgcgctt	ccggtgctca	ccgcgcgcca	6060
cgctcccgga	gcggtcgagt	tctggaccga	ccggctcggg	ttctcccggg	acttcgctga	6120
ggacgacttc	gcccgtgtgg	tcocggacga	cgtgacctg	ttcatcagcg	cggtccagga	6180
ccaggtggtg	ccggacaaca	ccctggcctg	ggtgtgggtg	cgccgctggg	acgagctgta	6240
cgccgagtgg	tcggaggtcg	tgtccacgaa	cttccgggac	gcctccgggc	cgcccatgac	6300
cgagatcggc	gagcagccgt	gggggocggg	gttcgcccctg	cgcgacctcg	ccggcaactg	6360
cgtgcacttc	gtggccgagg	agcaggactg	acacgtgcta	aaacttcatt	tttaatttaa	6420
aaggatctag	gtgaagatcc	tttttgataa	tctcatgacc	aaaatccctt	aacgtgagtt	6480
ttcgttccac	tgagcgtcag	accccgtaga	aaagatcaaa	ggatcttctt	gagatccttt	6540
ttttctgcgc	gtaatctgct	gcttgcaaac	aaaaaaccca	ccgctaccag	cggtgggttg	6600
tttgccggat	caagagctac	caactctttt	tccgaaggta	actggcttca	gcagagcgca	6660
gataccaaat	actgttcttc	tagttagacc	gtagttaggc	caccacttca	agaactctgt	6720
agcaccgctc	acatacctcg	ctctgcta	cctgttacca	gtggtgctg	ccagtggcga	6780
taagtcgtgt	cttaccgggt	tggactcaag	acgatagtta	ccgataagg	cgagcggctc	6840
gggctgaaac	gggggttcgt	gcacacagcc	cagcttgagg	cgaacgacct	acaccgaact	6900
gagataccta	cagcgtgagc	tatgagaaag	cgccacgctt	cccgaaggga	gaaagcggga	6960
caggtatccg	gtaagcggca	gggtcggaac	aggagagcgc	acgagggagc	ttccaggggg	7020
aaacgcctgg	tatctttata	gtcctgtcgg	gtttcgccac	ctctgacttg	agcgtcgatt	7080
tttgtgatgc	tcgtcagggg	ggcggagcct	atggaaaaac	gccagcaacg	cgcccttttt	7140
acggttctcg	gccttttget	ggccttttgc	tcacatgttc	tttctgctg	tatcccctga	7200
ttctgtggat	aaccgtatta	ccgcctttga	gtgagctgat	accgctcgcc	gcagccgaac	7260
gaccgagcgc	agcgagtcag	tgagcgagga	agcggagag	cgcccaatac	gcaaaccgcc	7320
tctccccgcg	cggtggccga	ttcattaatg	cagctggcac	gacaggtttc	ccgactggaa	7380
agcgggcagt	gagcgcaacg	caattaatgt	gagttagctc	actcattagg	cacccagggc	7440
tttacacttt	atgcttcocg	ctcgtatgtt	gtgtggaatt	gtgagcggat	aacaatttca	7500
cacaggaaac	agctatgacc	atgattacgc	caagctattt	aggtgacgcg	ttagaatact	7560

-continued

---

caagctatgc atcaagcttg gtaccgagct cggatccact agtaacggcc gccagtggtgc	7620
tggaattcag gggcgcgccc tatagatggg atgaagctgc tctcgacaaa tctgataaaa	7680
ctaaagaagg ttagtaatca atttttacaa aatcatagat tatttttttc attgaattat	7740
ttttatgcta taccaagaat tgtatttttag tatttgtttt aactacatat aatagaatta	7800
actacatata aattaactaa acttaaaata aaaatagatt tgmttcctga aattatttta	7860
agaatatata tgtatatatc taaaatctta gacttagata gatttttcta tctatctatt	7920
ttggttactt aaaataaata aatttgtata aataattgta tagttatcaa aaattaaaac	7980
taattttttt aaagttgttg atatataaaa tactaaagat ttaacgatta agtattttat	8040
taagtataga attttgtttt ttttttaagt ttagttatga agttgttaat tatattaaaa	8100
caaaacaata tttcgaaatt ttattatcat attcgaatat atttttttta gtgatgatgt	8160
atgaattatt atcataatth gaaagtttac taaaaaatat atcaacatga attgtaatat	8220
atgagttatt accttaacca aaattataaa ttaacattaa atataattat atagtccata	8280
tttagccata caatgtgtca tcaatattaa tagtcatgtc aatattacat aatgccata	8340
ttatgctact taaaccccaa atcccctaac tcccgtttaag tagccaaatt cataaatata	8400
cttattcgac aaaataaaaa actttaaaat atttactaat cggaccatgc acaagcatcc	8460
attccctatt ccattgccc gggataacaa tgcaaccnac tcctcaaaaa aagaaaaatt	8520
caagctcttt tgcaaaaaaa aataaaataa ttttaacacc taaaattttt tgmttccaaa	8580
cttctacagg gaacacacat aaaagaaaaa gaggacgtcc actcggatca cgcaacaaac	8640
caaaaggtgt gtcatgactc ctaagatata atatttcctt attcaaaatc ataccatttt	8700
aaattatgaa tgtatttctg agtccaccag atagttaatc caccagcgtt caaaccaag	8760
ttttatgatt gtaagtttaa gtgaattata ataatatatt cttaacggta tcttttcata	8820
actaattgag ttatcaaaact tgatcgaca tgtggctttg ataggtgtga cttttatggt	8880
atacaattct ttcaacctaa aaacattatt gttcctcaat atcttacatt atgcttgact	8940
gcaacaaaat atttttctcat ctgttttctt cctttaaacc aatttattat catctatttc	9000
ctgacatttt aatccatcca cctatgtcaa aaacttatag aaaatgtcaa cttccaaaca	9060
aaacataatt gaacttcgca aataaattct taataatatt aaaaaatgtt acttaattat	9120
ttcttcaacc ccattttcgg cgcgtagcgc ggacaaagac tctagttaaa tatagaagtt	9180
tccgattctc atcgtataaa acggtgactt tggcgggctt tcatgtgtaa caaattgggt	9240
taacaaaacca ctgcctagtc gtttagtgta gaatcagcgc atggaactcc gattggagcg	9300
tgactttcac gtgcgggagg cccaccacca cagcgggctt tacgctctaa gaatctcgcc	9360
cacggttttc ttcactctgcc ccccgccaag tgtcttcctc gttcggcact tctcaccag	9420
ttacaggaac cctaaaaatg gcctttcttc agccccgctt ataatacaca catgatccta	9480
tagtgggttc ttccacaagt tacatctcct tctggattgt acatttcaag tgtttgtgtt	9540
ttttctgcct ctgagagaaa atcgcggcgc catggctgct gctcccagtg tgaggacgtt	9600
tactcgggoc gaggttttga atgccgagc tctgaatgag ggcaagaagg atgccgagc	9660
acccttcttg atgatcatcg acaacaaggt gtacgatgtc cgcgagttcg tccctgatca	9720
tcccgggtga agtvtgatc tcacgcacgt tggcaaggac ggcactgacg tctttgacac	9780
ttttcaccoc gaggtctgct gggagactct tgccaacttt tacgttgggtg atattgacga	9840
gagcgaccgc gatatcaaga atgatgactt tgcggccgag gtccgcaagc tgcgtacctt	9900
gttccagtct cttggttact acgattcttc caaggcatat tacgccttca aggtctcgtt	9960

-continued

---

```

caacctctgc atctgggggtt tgtcgacggt cattgtggcc aagtggggcc agacctcgac 10020
cctcgccaac gtgctctcgg ctgcgctttt gggctctgtc tggcagcagt gcggatggtt 10080
ggctcacgac tttttgcac accaggctct ccaggaccgt ttctgggggtg atcttttcg 10140
cgccttcttg ggagggtgtc gccagggtct ctcgtcctcg tgggtggaagg acaagcacia 10200
cactcaccac gccgccccca acgtccacgg cgaggatccc gacattgaca cccaccctct 10260
gttgacctgg agtgagcatg cgttggagat gttctcggat gtcccagatg aggagctgac 10320
ccgcatgtgg tcgcgtttca tggctcgtgaa ccagacctgg ttttacttcc ccattctctc 10380
gtttgccctg ctctcctggt gcctccagtc cattctcttt gtgctgctta acggtcaggc 10440
ccacaagccc tcggggcgcgc gtgtgcccac ctcgttggtc gagcagctgt cgcttgcgat 10500
gcactggacc tggtaacctc ccaccatggt cctgttcac aaggatcccc tcaacatgct 10560
gggtgacttt ttggtgtcgc aggcgggtgt cggaacttg ttggcgatcg tgttctcgtc 10620
caaccacaac ggtatgcctg tgatctcgaa ggaggaggcg gtcgatatgg atttcttcac 10680
gaagcagatc atcacgggtc gtgatgtcca cccgggteta tttgccaact ggttcacggg 10740
tggattgaac tatcagatcg agcaccactt gttcccttcg atgcctcgc acaacttttc 10800
aaagatccag cctgctgtcg agaccctgtg caaaaagtac aatgtccgat accacaccac 10860
cggatgatc gagggaaactg cagaggctct tagccgtctg aacgaggctt ccaaggctgc 10920
ctccaagatg ggtaaggcgc agtaagc 10947

```

```

<210> SEQ ID NO 69
<211> LENGTH: 20742
<212> TYPE: DNA
<213> ORGANISM: Artificial sequence
<220> FEATURE:
<223> OTHER INFORMATION: Vector ARAL077
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: (14134)..(14134)
<223> OTHER INFORMATION: n is a, c, g, or t
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: (18932)..(18932)
<223> OTHER INFORMATION: n is a, c, g, or t
<400> SEQUENCE: 69

```

```

cgcgcctcga gtggggcgat cccccgggct gcaggaatc actggccgtc gttttacaac 60
gtcgtgactg ggaaaacctt ggcgttacc aacttaatcg ccttgacga catccccctt 120
tcgccagctg gcgtaatagc gaagaggccc gcaccgatcg cccttcccaa cagttgcgca 180
gcctgaatgg cgaatggatc gatccatcgc gatgtacctt ttgttagtca gcctctcgat 240
tgctcatcgt cattacacag taccgaagt ttgatgatct agtaacatag atgacaccgc 300
gcgcgataat ttatcctagt ttgcgcgcta tttttgttt tctatcgcgt attaaatgta 360
taattgcggg actctaata taaaaacca tctcataaat aacgtcatgc attacatggt 420
aattattaca tgcttaacgt aattcaacag aaattatatg ataactatcg caagaccggc 480
aacaggatcc aatcttaaga aactttattg ccaaatgttt gaacgatctg cttcgacgca 540
ctccttcttt actccaccat ctcgtcctta ttgaaaactg gggtagcacc aaaacgaatc 600
aagtcgctgg aactgaagtt accaatcacg ctggatgatt tgccagttgg attaactctg 660
cctttccccg catgaataat attgatgaat gcatgcgtga ggggtagtct gatgttgcca 720
atagctgcaa ttgccgcgac atcctccaac gagcataatt cttcagaaaa atagcgatgt 780

```

-continued

---

tccatgttgt cagggcatgc atgatgcacg ttatgaggtg acggtgctag gcagtattcc	840
ctcaaagttt catagtcagt atcatattca tcattgcatt cctgcaagag agaattgaga	900
cgcaatccac acgctgcggc aaccttccgg cgttcgtggt ctatttgctc ttggacgttg	960
caaacgtaag tgttgatcg atccggggtg ggcaagaac tccagcatga gatccccgcy	1020
ctggaggatc atccagccgg cgtcccggaa aacgattccg aagcccaacc tttcatagaa	1080
ggcggcggtg gaatcgaaat ctctgatgg caggttgggc gtcgcttggc cggtcatttc	1140
gaaccccaga gtcccgtca gaagaactcg tcaagaaggc gatagaaggc gatgctctgc	1200
gaatcgggag cggcgatacc gtaaagcacg aggaagcggc cagccattc gccgccaagc	1260
tcttcagcaa tateacgggt agccaacgct atgtctgat agcggctcgc cacaccagc	1320
cggccacagt cgatgaatcc agaaaagcgg ccattttcca ccatgatatt cggcaagcag	1380
gcatcgccat gggtaecgac gagatcctcg ccgtcgggca tgcgcgcctt gagcctggcg	1440
aacagtccgg ctggcgcgag cccctgatgc tcttcgtcca gatcatcctg atcgacaaga	1500
ccggcttcca tccgagtagc tgctcgtcgc atgctgatgt tcgcttgggt gtcgaatggg	1560
caggtagccg gatcaagcgt atgcagccgc cgcattgcat cagccatgat ggatactttc	1620
tcggcaggag caaggtgaga tgacaggaga tctgccccg gcacttcgcc caatagcagc	1680
cagtcccttc ccgcttcagt gacaacgctg agcacagctg cgcaaggaac gcccgctcgtg	1740
gccagccacg atagccgcgc tgccctcgtc tcaggttcat tcagggcacc ggacaggtcg	1800
gtcttgacaa aaagaaccgg gcgcccctgc gctgacagcc ggaacacggc ggcacagag	1860
cagccgattg tctgttctgc ccagtcatag ccgaatagcc tctccacca agcggccgga	1920
gaacctgcgt gcaatccatc ttgttcaatc atgcaaacg atccccgcaa gcttgagac	1980
tggtgatttc agcgtgtcct ctccaaatga aatgaacttc cttatataga ggaagggctc	2040
tgcaaggat agtgggattg tgcgtcatcc cttacgtcag tggagatac acatcaatcc	2100
acttgctttg aagacgtggt tggaacgtct tctttttcca cgatgctcct cgtgggtggg	2160
ggtccatctt tgggaccact gtcggcagag gcatcttcaa cgatggcctt tcctttatcg	2220
caatgatggc attttagga gccaccttc ttttccacta tcttcacaat aaagtgacag	2280
atagctgggc aatggaatcc gaggagggtt ccggatatta ccctttgttg aaaagtctca	2340
attgcccttt ggtcttctga gactgtatct ttgatatttt tggagtagac aagcgtgctg	2400
tgctccacca tgttgacgaa gattttcttc ttgtcattga gtcgtaagag actctgtatg	2460
aactgttcgc cagtctttac ggcgagttct gttaggctct ctatttgaat ctttgactcc	2520
atggcctttg attcagtggt aactacctt ttagagactc caatctctat tacttgctct	2580
ggtttgtgaa gcaagccttg aatcgtccat actggaatag tacttctgat cttgagaaat	2640
atatctttct ctgtgttctt gatgcagtta gtcctgaatc ttttgactgc atctttaacc	2700
ttcttgggaa ggtatttcat ctccctggaga ttattgctcg ggtagatcgt cttgatgaga	2760
cctgctgcgt aagcctctct aaccatctgt gggttagcat tctttctgaa attgaaaagg	2820
ctaactctct cattatcagt ggtgaacatg gtatcgtcac cttctccgct gaacttctg	2880
actagatcgt agagatagag gaagtcgtcc attgtgatct ctggggcaaa ggagatctga	2940
attaattcga tatggtggat ttatcacaaa tgggacccgc cgccgacaga ggtgtgatgt	3000
taggccagga ctttgaanaa ttgcgcaact atcgtatagt ggccgacaaa ttgacgccga	3060
gttgacagac tgcctagcat ttgagtgaat tatgtgaggt aatgggctac actgaattgg	3120
tagctcaaac tgtcagtatt tatgtatatg agtgtatatt ttccgataat ctgacocaa	3180

-continued

---

tctgaagatg	aaatgggtat	ctgggaatgg	cgaaatcaag	gcatcgatcg	tgaagtttct	3240
catctaagcc	cccatttga	cgtgaatgta	gacacgtcga	aataaagatt	tccgaattag	3300
aataatttgt	ttattgcttt	cgctataaa	tacgacggat	cgtaatttgt	cgttttatca	3360
aatgtactt	tcattttata	ataacgctgc	ggacatctac	atTTTTgaat	tgaaaaaaaa	3420
ttggtaatta	ctctttcttt	ttctccatat	tgaccatcat	actcattgct	gatccatgta	3480
gatttcccg	acatgaagcc	atttacaatt	gaatataatc	tgccgcccgt	gcccgtttgc	3540
accgggtgga	gcttgcattg	tggtttctac	gcagaactga	gcccgttagg	cagataattt	3600
ccattgagaa	ctgagccatg	tgacacttcc	ccccaacacg	gtgagcgacg	gggcaacgga	3660
gtgatccaca	tgggactttt	aaacatcatc	cgtcggatgg	cgttgcgaga	gaagcagtcg	3720
atccgtgaga	tcagccgacg	caccgggcag	gcgcgcaaca	cgatcgcaaa	gtatttgaac	3780
gcaggtacaa	tcgagccgac	gttcacgcgg	aacgaccaag	caagctagct	ttaatgcggg	3840
agtttatcac	agttaaattg	ctaacgcagt	caggcacccg	gtatgaaatc	taacaatgcg	3900
ctcatcgtea	tcctcggcac	cgtcacctg	gatgctgtag	gcataggctt	ggttatgccc	3960
gtactgccc	gcctcttgcg	ggatatcgtc	cattccgaca	gcatcgccag	tcactatggc	4020
gtgctgctag	cgctatatgc	gttgatgcaa	tttctatgcg	caccogttct	cggagcactg	4080
tccgaccgct	ttggcccgcg	cccagtcctg	ctcgettcgc	tacttggagc	cactatcgac	4140
tacgcgatca	tggcgaccac	accgcctctg	tggccaacc	cctccgctgc	tatagtgcag	4200
tcggcttctg	acgttcagtg	cagccgtctt	ctgaaaaaga	catgtcgcac	aagtcctaag	4260
ttacgcgaca	gggtgcgcgc	ctgccccttt	cctggcgttt	tcttgcgcg	tgttttagtc	4320
gcataaagta	gaataactgc	gactagaacc	ggagacatta	cgccatgaac	aagagcgcg	4380
ccgctggcct	gctgggctat	gcccgcgca	gcaccgacga	ccaggacttg	accaaccaac	4440
gggcccgaact	gcacgcggcc	gggtgcacca	agctgttttc	cgagaagatc	accggcacca	4500
ggcgcgacg	cccggagctg	gccaggatgc	ttgaccacct	acgcctggc	gacgttgga	4560
cagtgaccag	gctagaccgc	ctggcccgc	gcaccgcga	cctactggac	attgcccgagc	4620
gcatccagga	ggcccggcgc	ggcctgcgta	gcccggcaga	gcccgtgggc	gacaccacca	4680
cgcccggccg	ccgcatgggt	ttgaccgtgt	tcgcccggat	tgccgagttc	gagcgttccc	4740
taatcatcga	ccgcaccgcg	agcgggccc	aggcccga	ggcccagggc	gtgaagtttg	4800
gccccgcgcc	tacctcacc	ccggcacaga	tcgcccgcgc	ccgcgagctg	atcgaccagg	4860
aaggccgcac	cgtgaaagag	gcggctgcac	tgcttggcgt	gcatcgctcg	accctgtacc	4920
gcgcacttga	gcgcagcgag	gaagtgcgc	ccaccgagc	caggcggcgc	ggtgccttcc	4980
gtgaggacgc	attgaccgag	gcccagccc	tggcggccc	cgagaatgaa	cgccaagagg	5040
aacaagcatg	aaaccgcacc	aggacggcca	ggacgaaccg	ttttcatta	ccgaagagat	5100
cgaggcggag	atgatcgcg	ccgggtacgt	gttcgagccg	cccgcgcacg	tctcaaccgt	5160
gcggctgcat	gaaatcctgg	ccggtttgtc	tgatgccaag	ctggcggcct	ggcggccag	5220
cttggccgct	gaagaaaccg	agcgcgcgcg	tctaaaaagg	tgatgtgtat	ttgagtaaaa	5280
cagcttgcgt	catcgggtcg	ctcgttatat	gatgcgatga	gtaaataaac	aaatacgc	5340
gggaacgc	gaagttatcg	ctgtacttaa	ccagaaaggc	gggtcaggca	agacgaccat	5400
cgcaaccat	ctagcccgcg	ccctgcaact	cgcccgggccc	gatgttctgt	tagtcgattc	5460
cgatccccag	ggcagtgccc	gcgattgggc	ggcgcgtgcg	gaagateaac	cgctaaccgt	5520

-continued

---

tgtcggcatc gaccgcccga cgattgaccg cgacgtgaag gccatcggcc ggcgcgactt	5580
cgtagtgatc gacggagcgc cccaggcggc ggacttggt gtgtccgcga tcaaggcagc	5640
cgacttcgtg ctgattccgg tgcagccaag cccttacgac atatgggcca ccgccgacct	5700
ggtggagctg gttaagcagc gcattgaggt cacggatgga aggctacaag cggcctttgt	5760
cgtgtcgcgg gcgatcaaa gcaacgcgat cggcggtag gttgccgagg cgctggccgg	5820
gtacgagctg ccattcttg agtcccgat cacgcagcgc gtgagctacc caggcactgc	5880
cgcccccggc acaaccgttc ttgaatcaga acccgagggc gacgctgccc gcgaggtcca	5940
ggcgtggcc gctgaaatta aatcaaaact catttgagtt aatgaggtaa agagaaaatg	6000
agcaaaagca caaacacgct aagtgcggc cgtccgagcg cacgcagcag caaggctgca	6060
acgttgcca gcctggcaga cacgccagcc atgaagcggg tcaactttca gttgccggcg	6120
gaggatcaca ccaagctgaa gatgtacgcg gtacgccaa gcaagaccat taccgagctg	6180
ctatctgaat acatcgcgca gctaccagag taaatgagca aatgaataaa tgagtagatg	6240
aattttagcg gctaaaggag gcggcatgga aaatcaagaa caaccaggca ccgacgccgt	6300
ggaatgcccc atgtgtggag gaacggggcg ttggccaggc gtaagcggct gggttgtctg	6360
ccggccctgc aatggcactg gaacccccaa gcccgaggaa tcggcgtgag cggtcgcaaa	6420
ccatccggcc cggtagcaaat cggcggggcg ctgggtgatg acctggtgga gaagttgaag	6480
gccgcgcagg ccgccagcg gcaacgcata gaggcagaag cacgcccgg tgaatcgtgg	6540
caagcggccg ctgatcgaat ccgcaaaagaa tcccggcaac cgccggcagc cggtgccgcc	6600
tcgattagga agccgcccaa gggcgacgag caaccagatt ttttcgttcc gatgctctat	6660
gacgtgggca ccccgatag tcgcagcata atggacgtgg ccgttttccg tctgtcgaag	6720
cgtgaccgac gagctggcga ggtgatccgc tacgagcttc cagacgggca cgtagaggtt	6780
tccgcagggc cggccggcat ggccagtgtg tgggattacg acctggtact gatggcggtt	6840
tcccactaa ccgaatccat gaaccgatac cgggaaggga agggagacaa gcccgcccg	6900
gtgttccgtc cacacgttgc ggacgtactc aagttctgcc ggcgagccga tggcggaaa	6960
cagaaagacg acctggtaga aacctgcatt cggttaaaca ccacgcacgt tgccatgcag	7020
cgtacgaaga aggccaaagaa cggccgcctg gtgacgggat ccgagggatga agccttgatt	7080
agccgctaca agatcgtaaa gagcgaaacc gggcggccgg agtacatcga gatcgagcta	7140
gctgattgga tgtaccgca gatcacagaa ggcaagaacc cggacgtgct gacggtcac	7200
cccgattact ttttgatcga tcccggcata gcccgtttcc tctaccgctt gccacgccgc	7260
gccgcaggca aggcagaagc cagatggttg ttcaagacga tctacgaacg cagtggcagc	7320
gccggagagt tcaagaagtt ctgtttcacc gtgcgcaagc tgatcgggtc aatgacctg	7380
ccggagtacg atttgaagga ggaggcgggg caggctggcc cgatcctagt catgcgctac	7440
cgcaacctga tcgagggcga agcatccgcc ggttcctaata gtacggagca gatgctaggg	7500
caaattgccc tagcagggga aaaaggtcga aaaggtctct ttctgtgga tagcacgtac	7560
attgggaacc caaagccgta cattgggaac cggaacccgt acattgggaa cccaaagccg	7620
tacattggga accggtcaca catgtaagtg actgatataa aagagaaaaa aggcgatttt	7680
tccgcctaaa actctttaa acttattaaa actcttaaaa cccgcctggc ctgtgcataa	7740
ctgtctggcc agcgcacagc cgaagagctg caaaaagcgc ctacccttcg gtcgctgcgc	7800
tccctacgcc ccgccgcttc gcgtcggcct atccggggcg ctggccgctc aaaaatggct	7860
ggcctacggc caggcaatct accagggcgc ggacaagccg ccgcgctgcc actcgaccgc	7920

-continued

---

cggcgcccac	atcaaggcac	cctgcctcgc	gcgtttcggg	gatgacggg	aaaacctctg	7980
acacatgcag	ctccccgaga	cggtcacagc	ttgtctgtaa	gcggatgccg	ggagcagaca	8040
agccccgcag	ggcgcgcag	cgggtgttg	cgggtgtcgg	ggcgcagcca	tgacctcagtc	8100
acgtagcgat	agcggagtg	atactggctt	aactatgcgg	catcagagca	gattgtactg	8160
agagtgcacc	atatgcggg	tgaataaccg	cacagatgcg	taaggagaaa	ataccgcatc	8220
aggcgctctt	ccgcttctc	gctcaactgac	tcgctgcgct	cggtcgttcg	gctgcggcga	8280
gcggtatcag	ctcactcaaa	ggcggtaata	cggttatcca	cagaatcagg	ggataacgca	8340
ggaaagaaca	tgtgagcaaa	aggccagcaa	aaggccagga	accgtaaaaa	ggcgcggtg	8400
ctggcgtttt	tccataggtc	ccgccccct	gacgagcatc	acaaaaatcg	acgctcaagt	8460
cagaggtggc	gaaacctgac	aggactataa	agataaccag	cgtttcccc	tggaagctcc	8520
ctcgtgcgct	ctcctgttcc	gacctgcgcg	cttacgggat	acctgtccgc	ctttctccct	8580
tcgggaagcg	tggcgcttcc	tcatagctca	cgctgtaggt	atctcagttc	ggtgtaggtc	8640
gttcgctcca	agctgggctg	tgtgcacgaa	cccccgcttc	agccccagccg	ctgcgcctta	8700
tccggttaact	atcgtcttga	gtccaacctg	gtaagacacg	acttatcgcc	actggcagca	8760
gccactggta	acaggattag	cagagcgagg	tatgtaggcg	gtgctacaga	gttcttgaag	8820
tggtggccta	actacggcta	cactagaagg	acagtatttg	gtatctgcgc	tctgctgaag	8880
ccagttacct	tcggaaaaag	agttggtagc	tcttgatccg	gcaaaaaaac	caccgctggt	8940
agcgttggtt	tttttgtttg	caagcagcag	attacgcgca	gaaaaaaagg	atctcaagaa	9000
gatcctttga	tcttttctac	ggggctctgac	gctcagtgga	acgaaaactc	acgttaaggg	9060
atthttgtca	tgagattatc	aaaaaggatc	ttcacctaga	tccttttaaa	ttaaaaatga	9120
agtttttaaat	caatctaaag	tatatatgag	taaacttggt	ctgacagtta	ccaatgctta	9180
atcagtgagg	cacctatctc	agcgatctgt	ctatttctgt	catccatagt	tgctgactc	9240
cccgtcgtgt	agataactac	gatacgggag	ggcttaccat	ctggccccag	tgctgcaatg	9300
ataccgcgag	accacgcctc	accggctcca	gatttatcag	caataaacca	gccagccgga	9360
agggccgagc	gcagaagtgg	tctgcaact	ttatccgcct	ccatccagtc	tattaattgt	9420
tgccgggaag	ctagagtaag	tagttcgcca	gtaaatagtt	tgcgcaactg	tgttgccatt	9480
gctacaggca	tcgtgggtgc	acgctcgtcg	tttggtatgg	cttcattcag	ctccggttcc	9540
caacgatcaa	ggcaggttac	atgatcccc	atggtgtgca	aaaaagcggg	tagctccttc	9600
ggtcctccga	tcgttgtcag	aagtaagttg	gccgcagtg	tatcactcat	ggttatggca	9660
gcactgcata	attctcttac	tgtcatgcca	tccgtaagat	gcttttctgt	gactggtgag	9720
tactcaacca	agtcattctg	agaatagtg	atgcggcgac	cgagttgctc	ttgccccgcg	9780
tcaacacggg	ataataccgc	gccacatagc	agaactttaa	aagtgctcat	cattggaaaa	9840
gacctgcagg	gggggggggg	cgctgaggtc	tgctcctgga	agaaggtggt	gctgactcat	9900
accaggcctg	aatcgcccc	tcattccagcc	agaaagtgag	ggagccacgg	ttgatgagag	9960
ctttgttgta	ggtggaccag	ttggtgattt	tgaacttttg	ctttgccacg	gaacgggtctg	10020
cgttgtcggg	aagatgcgtg	atctgatcct	tcaactcagc	aaaagttcga	tttattcaac	10080
aaagccgccc	tcccgtcaag	tcagcgtaat	gctctgccag	tggtacaacc	aattaacca	10140
ttctgattag	aaaaactcat	cgagcatcaa	atgaaactgc	aatttattca	tatcaggatt	10200
atcaatacca	tatthttgaa	aaagccggtt	ctgtaatgaa	ggagaaaact	caccgaggca	10260

-continued

---

gttccatagg atggcaagat cctggtatcg gtctgcgatt ccgactcgtc caacatcaat 10320  
 acaacctatt aatttcccct cgtaaaaaat aaggttatca agtgagaaat caccatgagt 10380  
 gacgactgaa tccggtgaga atggcaaaag cttatgcatt tctttccaga cttggtcaac 10440  
 aggccagcca ttacgctcgt catcaaaatc actcgcacat accaaaccgt tattcattcg 10500  
 tgattgcgcc tgagcgagac gaaatacgcg atcgcgtgta aaaggacaat tacaaacagg 10560  
 aatcgaatgc aaccggcgca ggaacactgc cagcgcacat acaatatttt cacctgaatc 10620  
 aggatattct tctaatacct ggaatgctgt tttcccgggg atcgcagtgg tgagtaacca 10680  
 tgcacatca ggagtacgga taaatgctt gatggtcggg agaggcataa attccgtcag 10740  
 ccagtttagt ctgaccatct catctgtaac atcattggca acgctacctt tgccatgttt 10800  
 cagaacaac tctggcgcac cgggcttccc atacaatcga tagattgtcg cacctgattg 10860  
 cccgacatta tcgagagccc atttataccc atataaatca gcatccatgt tggaatttaa 10920  
 tcgcgccctc gagcaagacg tttcccgttg aatatggctc ataacacccc ttgtattact 10980  
 gtttatgtaa gcagacagtt ttattgttca tgatgatata tttttatctt gtgcaatgta 11040  
 acatcagaga ttttgagaca caactggctc tcccccccc cccctgcagg tcaattcgggt 11100  
 cgatatggct attacgaaga aggctcgtgc gcggagtcgc gtgaacttcc ccacgcaaca 11160  
 agtgaaccgc accgggtttg ccggaggcca tttcgttaaa atgcgcagcc atggctgctt 11220  
 cgtccagcat ggcgtaatac tgatcctcgt cttcggctgg cggatatattg ccgatgggct 11280  
 tcaaaagccg ccgtggttga accagtcctat ccattccaag gtagcgaact cgaccgcttc 11340  
 gaagctcttc catggtccac gccgatgaat gacctcggcc ttgtaaagac cgttgatcgc 11400  
 ttctcgcagg gcgttgctgt gctgtcgcgc acgcttccga tagatggctc gatacctgct 11460  
 tctgccaacc gctcgggaata gcgaaaggac acgtattgaa caccgcgacg cgagtgatgc 11520  
 actaggccgc catgagcggg acgccgatca tgatgagcct cctcagagggc atcagaggaca 11580  
 aagcctgcat gtgctgtccg gctcgcgccg catccgacaa tgcgacgggc gaagacgtcg 11640  
 atcacgaagg ccacgtagac gaagccctcc caagtggcga cataagtacg gacatgcgca 11700  
 aaggctttcc cggtttgcgc ctgatgggtc aagagacgct gaagcgcgat ccgatgcgca 11760  
 ggcactctgt cgtcttcgcg ggtcgtggcg gtggcctgat caaggctcact cgccgaagag 11820  
 ctgcatgatt ggtcgcgaac cgagcggggg aaattgtcgc gcagttctcc cgtcgcgag 11880  
 gcgataaatt acatgctcaa gcgatgggat ggcattacgt cattcctcga tgacggcccg 11940  
 atttgcctga cgaacaatgc tgccgaacga acgctcagag gctatgtact cggcaggaag 12000  
 tcatggctgt ttgcccgatc ggatcgttgt gctgaacgtg cggcgttcat ggcgacactg 12060  
 atcatgagcg ccaagctcaa taacatcgat ccgcaggcct ggcttgccga cgtcccgcgc 12120  
 gaccttgccg acgctccgat cagcaggctt gagcaacagc tgccgtggaa ctggacatcc 12180  
 aagacactga gtgctcaggc ggctgacct gcggccttca ccggatactt accccattat 12240  
 cgcagattgc gatgaagcat cagcgtcatt cagcaatctt gccaaagtat gcaggctcgc 12300  
 gagaatcgac gtgcgaaacc ggtcgtgtgc gccaaagatc cgcttgccga gcggtcgaac 12360  
 attcatgctg ggacttcaag aggtcagta gaggaagaac cggaagggtt gcaccggaaa 12420  
 atatgcgttc ctttgagagc gcctcatgg acgtgaacaa atcgcgccga ccaaggatgc 12480  
 cacggataca aaagctcgcg aagctcggtc ccgtgggtgt tctgtcgtct cgttgtacaa 12540  
 cgaaatccat tcccattccg cgctcaagat ggcttcccct cggcagttca tcagggctaa 12600  
 atcaatctag ccgacttgtc cgtgtaaatg ggctgcactc caacagaaac aatcaaaaaa 12660

-continued

---

acatacacag cgacttatcc acacgagctc aaattacaac ggtatatatc ctgccagtca 12720  
gcatcatcac accaaaagtt aggcccgat agtttgaat tagaaagctc gcaattgagg 12780  
tctacaggcc aaattcgctc ttagccgtac aatattactc accgggtgca tgcccccat 12840  
cgtaggtgaa ggtggaaatt aatgatccat cttgagacca caggcccaca acagctacca 12900  
gtttcctcaa gggccacca aaaacgtaag cgcttacgta catggctgat aagaaaaggc 12960  
aatttgtaga tgtaacatc caacgtcgtc ttcagggatc gatccaatac gcaaaccgcc 13020  
tctccccgcg cgttggccga ttcattaatg cagctggcac gacaggtttc ccgactggaa 13080  
agcgggcagt gagcgcaacg caattaatgt gagttagctc actcattagg caccocaggc 13140  
tttaccttt atgcttcogc ctgctatggt gtgtggaatt gtgagcggat aacaatttca 13200  
cacaggaaac agctatgacc atgattacgc caagcttgcg tgccctgcagg tcgactctag 13260  
aggatctggc gcgccctata gatgggatga agctgctctc gacaaatctg ataaaactaa 13320  
agaagggttag taatcaattt ttcaaaaac atagattatt tttttcattg aattattttt 13380  
atgctatacc aagaattgta ttttagtatt tgttttaact acatataata gaattaacta 13440  
catataaatt aactaaactt aaaataaaaa tagatttgtt tcctgaaatt attttaagaa 13500  
tatatatgta tatatctaaa atcttagact tagatagatt tttctatcta tctattttgg 13560  
ttacttaaaa taaataaatt tgtataaata attgtatagt tatcaaaaat taaaactaat 13620  
ttttttaaag ttgttgat ataaaact aaagatttaa cgattaagta tttatttaag 13680  
tatagaattt tgtttttttt ttaagtttag ttatgaagtt gttaattata ttaaaacaaa 13740  
acaatatttc gaaattttat tatcatatc gaatatattt tttttagtga tgatgtatga 13800  
attattatca taatttgaaa gtttactaaa aaatatatca acatgaattg taatatatga 13860  
gttattacct taacaaaat tataaattaa cattaatat aattatata gtcataatca 13920  
gccatacaat gtgtcatcaa tattaatagt catgtcaata ttacataatg ccaatattat 13980  
gctacttaaa ccccaaatcc cctaactccc gtttaagtagc caaattcata aatatactta 14040  
ttcgacaaaa taaaaaactt taaaatattt actaatccga ccatgcacaa gcatccattc 14100  
cctattccat tggccacggga taacaatgca accnactcct caaaaaaaga aaaattcaag 14160  
ctcttttgca aaaaaaata aaataattt aacacctaaa attttttgtt tccaaacttc 14220  
tacagggaac acacataaaa gaaaaagagg acgtccactc ggatcacgca acaaaccaaa 14280  
agggtgtgca tgactcctaa gatataatat ttccttattc aaaatcatac ctttttaaat 14340  
tatgaatgta tttcgtagtc caccagatat gtaatccacc agcgttcaaa ccaaagtttt 14400  
atgattgtaa gtttaagtga attataataa tatattcttc acggtatcct ttcataacta 14460  
attgagttat caaacttgat cgcacatgtg gctttgatag gtgtgacttt tatggtatac 14520  
aattctttca acctaaaaac attattgttc ctcaatatct tacattatgc ttgactgcaa 14580  
caaaaatatt tctcatctgt tttcttctt taaaccaatt tattatcact ttttctctga 14640  
cattttaatc catccaccta tgtcaaaaac ttatagaaaa tgtcaacttc caaacaacaa 14700  
ataattgaac ttgcgaaata aattcttaat aatattaaaa atgttactt aattatttct 14760  
tcaaccccat tttccgctgc tagcgcggac aaagactcta gttaaatata gaagtttccg 14820  
attctcatcg tataaacgg tgactttggc gggctttcat gtgtaacaaa ttggtttaa 14880  
aaaccactgc ctagtctgtt agtgtagaat cagcgcattg aactccgatt ggagcgtgac 14940  
tttcacgtgc cggaggccca ccaccacgc gggcgttacg ctctaagaat ctgcgccacg 15000

-continued

---

gtttttctca	tctgcccccc	gccaagtgtc	ttctctgttc	gccacttctc	accaagttac	15060
aggaacccta	aaaatggcct	ttcttcagcc	ccggctataa	tacacacatg	atcctatagt	15120
gggttcttcc	acaagttaca	tctccttctg	gattgtacat	ttcaagtgtt	tgtgtttttt	15180
ctgcctctga	gagaaaatcg	cgcccgcatg	gctgctgtc	ccagtgtgag	gacgtttact	15240
cgggccgagg	ttttgaatgc	cgaggctctg	aatgagggca	agaaggatgc	cgaggcacc	15300
ttcttgatga	tcctcgacaa	caagggtgac	gatgtccg	agttcgtccc	tgatcatccc	15360
ggtggaagtg	tgattctcac	gcacgttggc	aaggacggca	ctgacgtctt	tgacactttt	15420
caccccgagg	ctgcttggga	gactcttgcc	aacttttacg	ttggtgatat	tgacgagagc	15480
gaccgcgata	tcaagaatga	tgactttg	gccgaggtcc	gcaagctg	tacctgtt	15540
cagtctcttg	gttactacga	ttcttccaag	gcatactacg	ccttcaaggt	ctcgttcaac	15600
ctctgcatct	ggggtttgtc	gacggtcatt	gtggccaagt	ggggccagac	ctcgaccctc	15660
gccaactgtc	tctcggctgc	gcttttgggt	ctgttctggc	agcagtgcgg	atgggtggct	15720
cacgactttt	tgcatcacca	ggtcttccag	gaccgtttct	ggggtgatct	tttcggcgcc	15780
ttcttgggag	gtgtctgcc	gggcttctcg	tctctgtggt	ggaaggacaa	gcacaacact	15840
caccacgccc	cccccaactg	ccacggcgag	gateccgaca	tgacaccca	ccctctgttg	15900
acctggagtg	agcatgcggt	ggagatgttc	tcggatgtcc	cagatgagga	gctgaccg	15960
atgtggctgc	gtttcatggt	cctgaaccag	acctggtttt	acttcccct	tctctcgttt	16020
gccccctct	cctgggtgct	ccagtcatt	ctctttgtgc	tgctaacgg	tcaggcccac	16080
aagccctcgg	gcgcgctgt	gcccactctg	ttggtcgagc	agctgtcgt	tgatgac	16140
tggacctggt	acctcgccac	catgttctcg	ttcatcaagg	atcccgtaa	catgctggtg	16200
tactttttgg	tgctgcaggc	ggtgtcgga	aacttgttgg	cgatcgtgtt	ctcgtcaac	16260
cacaacggta	tgctgtgat	ctcgaaggag	gaggcggctg	atatggattt	cttcacgaag	16320
cagatcatca	cgggtcgtga	tgctccaccg	ggtctatttg	ccaactggtt	cacgggtgga	16380
ttgaactatc	agatcgagca	ccactgttcc	ccttcgatgc	ctcgccacaa	cttttcaaa	16440
atccagcctg	ctgtcgagac	cctgtgcaaa	aagtacaatg	tccgatacca	caccaccggt	16500
atgatcgagg	gaactgcaga	ggtctttagc	cgtctgaacg	aggtctccaa	ggctgcctcc	16560
aagatgggta	aggcgcagta	agcggccgca	agtatgaact	aaaatgcatg	taggtgtaag	16620
agctcatgga	gagcatggaa	tattgtatcc	gaccatgtaa	cagtataata	actgagctcc	16680
atctcacttc	ttctatgaat	aaacaaagga	tgttatgata	tattaacact	ctatctatgc	16740
accttattgt	tctatgataa	atttctctt	attattataa	atcatctgaa	tcgtgacggc	16800
ttatggaatg	cttcaaatag	tacaaaaaca	aatgtgtact	ataagacttt	ctaaacaatt	16860
ctaaccttag	cattgtgaac	gagacataag	tgtaagaag	acataacaat	tataatggaa	16920
gaagtttgtc	tccatttata	tattatata	taccactta	tgtattat	taggatgta	16980
aggagacata	acaattataa	agagagaagt	ttgatccat	ttatatatta	tatactacc	17040
atztatatat	tatacttata	cacttattta	atgtctttat	aaggtttgat	ccatgatatt	17100
tctaataatt	tagttgat	gtatatgaaa	gggtactatt	tgaactctct	tactctgtat	17160
aaaggttgg	tcctccttaa	agtggtctta	tttaatttta	ttgcttctta	cagataaaaa	17220
aaaaattatg	agttggtttg	ataaaatatt	gaaggattta	aaataataat	aaataacata	17280
taatatatgt	ataaaatatt	attataatat	aacatttata	tataaaaaag	taaatattgt	17340
cataaatcta	tacaatcgtt	tagccttctg	ggacgaatct	caattattta	aacgagagta	17400

-continued

---

```

aacatatttg actttttggt tatttaacaa attattattt aacctatat gaaatTTTT 17460
ttttatcag caaagaataa aattaaata agaaggacaa tgggtgccca atccttatac 17520
aaccaacttc cacaagaag tcaagtcaga gacaacaaaa aaacaagcaa aggaaatTTT 17580
ttaatttgag ttgtcttgtt tgctgcataa tttatgcagt aaaacactac acataaccct 17640
tttagcagta gagcaatggt tgaccgtgtg cttagcttct tttatTTTat tttttatca 17700
gcaaagaata aataaaataa aatgagacac ttcagggatg tttcaacaag cttggatcct 17760
cgaagagaag ggtaataaac acactTTTTt aacatTTTTa acacaaatTT tagttatTTa 17820
aaaatttatt aaaaaatTTa aaataagaag aggaactcTT taaataaatc taacttaca 17880
aatTTatgat tTTtaataag tTTtcacaa taaaaaatgt cataaaaaata tgttaaaaag 17940
tatattatca atattctcTT tatgataaat aaaaagaaaa aaaaaataa agttaagtga 18000
aatgagatt gaagtgactt taggtgtgta taaatatatc aaccccgcca acaatttatt 18060
taatccaaat atattgaagt atattattcc atagcctTTa tttatTTata tatttattat 18120
ataaaagctt tatttgttct aggtgttca tgaaatattt ttttggTTTT atctccgTTg 18180
taagaaaatc atgtgctTTg tgcgccoact cactattgca gctTTTTcat gcattggtca 18240
gattgacggT tgattgtatt tttgtTTTTt atggTTTTgt gttatgactt aagtcttcat 18300
ctctttatct cttcatcagg tttgatggtt acctaatatg gtccatgggt acatgcattg 18360
ttaaattagg tggccaactt tgttgtgaac gatagaattt tttttatatt aagtaaaacta 18420
tttttatatt atgaaataat aataaaaaaa atattttatc attattaaca aaatcatatt 18480
agttaatttg ttaactctat aataaaagaa atactgtaac attcacatta catggttaaca 18540
tctttccacc ctttcatttg tttttgttt gatgactttt tttcttgttt aaatttattt 18600
cccttctttt aaatttggaa tacattatca tcatatataa actaaaaatc taaaaacagg 18660
attacacaaa tgataaataa taacacaaat atttataaat ctagctgcaa tatattttaa 18720
ctagctatat cgatattgta aaataaaact agctgcattg atactgataa aaaaatatca 18780
tgtgctttct ggactgatga tgcagtatac ttttgacatt gcctttattt tatttttcag 18840
aaaagctttc ttagtcttgg gttcttcatt atttgtttcc catctccatt gtgaattgaa 18900
tcatttgctt cgtgtcacia atacaattta gntaggtaca tgcattggtc agattcacgg 18960
tttattatgt catgacttaa gttcatggta gtacattacc tgccacgcat gcattatatt 19020
ggttagattt gataggcaaa tttggttgc aacaataa atataaataa tgtttttata 19080
ttacgaaata acagtgatca aaacaacag ttttatcttt attaacaaga tttgttttt 19140
gtttgatgac gttttttaaT gtttacgctt tcccccttct tttgaattta gaacactTTa 19200
tcatcataaa atcaaatact aaaaaaatta catatttcat aaataataac acaaatattt 19260
ttaaaaaatc tgaataata atgaacaata ttacatatta tcacgaaaat tcatataa 19320
aaatattata taaataaaat gtaatagtag ttatatgtag gaaaaaagta ctgcacgcat 19380
aatatataca aaaagattaa aatgaactat tataaataat aacctaaat taatggtgaa 19440
tcatatcaaa ataatgaaaa agtaaaaaa atttgtaatt aacttctata tgtattacac 19500
acacaaataa taaataatag taaaaaaat tatgataaat atttaccatc tcataagata 19560
tttaaaataa tgataaaaat atagattatt ttttatgcaa ctagctagcc aaaaagagaa 19620
cacgggtata tataaaaaa gtacctttaa attctactgt acttcttTTa ttctgacgt 19680
ttttatatca agtggacata cgtgaagatt ttaattatca gtctaaatat ttcattagca 19740

```

-continued

---

```

cttaatactt ttctgtttta ttcctatcct ataagtagtc ccgattctcc caacattgct 19800
tattcacaca actaactaag aaagtcttcc atagccccc aagcggccca tggcctcctc 19860
cgaggacgtc atcaaggagt tcatgcgctt caaggtgctc atggagggct cegtgaacgg 19920
ccacgagttc gagatcgagg gcgagggcga gggccgcccc tacgagggca cccagaccgc 19980
caagctgaag gtgaccaagg gcggccccct gcccttcgcc tgggacatcc tgtcccccca 20040
gttccagtac ggtccaagg tgtacgtgaa gcaccccgcc gacatccccg actacaagaa 20100
gctgtccttc cccgagggct tcaagtggga gcgctgatg aacttcgagg acggcggcgt 20160
ggtgaccgtg acccaggact cctccctgca ggacggctcc ttcactaca aggtgaagt 20220
catcggcgtg aacttccct cgcagggccc cgtaatgcag aagaagacta tgggctggga 20280
ggcctccacc gagcgcctgt acccccgcga cggcgtgctg aagggcgaga tccacaaggc 20340
cctgaagctg aaggacggcg gccactacct ggtggagttc aagtccatct acatggccaa 20400
gaagcccgtg cagctgcccg gctactacta cgtggactcc aagctggaca tcacctccca 20460
caacgaggac tacaccatcg tggagcagta cgagcgcgcc gagggcgcgc accacctgtt 20520
cctgtagcgg ccggccgcga cacaagtgtg agagtactaa ataatgctt tggttgtacg 20580
aatcattac actaataaaa ataataaag cttatatatg ccttcgcta aggccgaatg 20640
caaagaaatt ggttctttct cgttatcttt tgccactttt actagtagct attaattact 20700
acttaatcat ctttgtttac ggctcattat atccgtcgac gg 20742

```

```

<210> SEQ ID NO 70
<211> LENGTH: 10758
<212> TYPE: DNA
<213> ORGANISM: Artificial sequence
<220> FEATURE:
<223> OTHER INFORMATION: Vector KS431
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: (2349)..(2349)
<223> OTHER INFORMATION: n is a, c, g, or t
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: (8498)..(8498)
<223> OTHER INFORMATION: n is a, c, g, or t

<400> SEQUENCE: 70

```

```

ggccgcaagt atgaactaaa atgcatgtag gtgtaagagc tcatggagag catggaatat 60
tgtatccgac catgtaacag tataataact gagctccatc tcacttcttc tatgaataaa 120
caaaggatgt tatgatatat taacctcta tctatgcacc ttattgttct atgataaatt 180
tcctcttatt attataaatc atctgaatcg tgacggctta tggaaatgctt caaatagtag 240
aaaaacaaat gtgtactata agactttcta aacaattcta accttagcat tgtgaacgag 300
acataagtgt taagaagaca taacaattat aatggaagaa gtttgcctcc atttatatat 360
tatatattac ccacttatgt attatattag gatgttaagg agacataaca attataaaga 420
gagaagtttg tatccattta tatattatat actaccatt tatatattat acttatccac 480
ttatttaatg tctttataag gtttgatcca tgatatttct aatattttag ttgatatgta 540
tatgaaaggg tactatttga actctcttac tctgtataaa ggttgatca tccttaaagt 600
gggtctatct aattttattg cttcttacag ataaaaaaaa aattatgagt tggtttgata 660
aaatattgaa ggatttaaaa taataataaa taacataaa tatatgtata taaatttatt 720
ataatataac atttatctat aaaaaagtaa atattgtcat aaatctatac aatcgtttag 780
ccttgctgga cgaatctcaa ttatttaaac gagagtaaac atatttgact ttttggttat 840

```

-continued

---

ttaacaaatt	attatntaac	actatatgaa	atntntntnt	ttatcagcaa	agaataaaat	900
taaattaaga	aggacaatgg	tgtccaate	cttatacaac	caactccac	aagaaagtca	960
agtcagagac	aacaaaaaa	caagcaagg	aaatntntta	atntgagttg	tcttgnttg	1020
tgcataatnt	atgcagtaaa	acactacaca	taacctntt	agcagtagag	caatggntga	1080
ccgtgtgctt	agcttctntt	atnttatntt	tttatcagca	aagaataaa	aaaataaaat	1140
gagacacttc	agggatgntt	caacaagctt	ggatcctcga	agagaagggt	taataacaca	1200
ctntntntaac	atntntntaac	caaatnttag	ttatntntaa	atnttatntaa	aaatntntaa	1260
taagaagagg	aactctntta	ataaatctaa	cttacaaaat	ttatgatntt	taataagntt	1320
tcaccaataa	aaaatgtcat	aaaaatatgt	taaaaagtat	attatcaata	ttctctntt	1380
gataaataaa	aagaaaaaa	aaataaaagt	taagtgaaaa	tgagattgaa	gtgactnttag	1440
gtgtgtataa	atatatcaac	cccgcaca	atnttatntta	tccaaatata	ttgaagtata	1500
ttatccata	gcctntntt	atnttatata	ttatntntata	aaagctntt	ttgnttctag	1560
ttgntcatga	aatatntntt	tgntntntatc	tccgntgtaa	gaaaatcatg	tgctnttgnt	1620
cgccactcac	tattgcagct	ttntcatgca	ttgntcagat	tgacgnttga	ttgatntntt	1680
gntntntatg	gntntgntt	atgactntaag	tctntcatctc	ttatctctt	catcagntt	1740
gatgnttacc	taatatgntc	catgnttaca	tgcatgntta	aatntgnttg	ccaactntt	1800
tgntaacgat	agaatntntt	ttatntntaag	taactatntt	ttatntntatg	aaataaat	1860
aaaaaaata	ttnttatcatt	attaacaaaa	tcatntntag	taatntgntta	actctataat	1920
aaaagaaata	ctgntaacatt	cacntntacat	ggntaacatct	ttccacctt	tcatntgntt	1980
ttgntntgat	gactntntntt	ctgntntntaa	ttatntntccc	ttctntntaa	ttgntntaac	2040
attatcatca	tataaaact	aaaactntaa	aaacagntt	acacaaatga	taaaataaa	2100
cacaaatatt	tataaatcta	gctgcaatat	atntntaaact	gctatntcga	tattgntaaaa	2160
taaaactagc	tgcatntgata	ctgataaaaa	aatatcatgt	gctntctgga	ctgatgatgc	2220
agtatactntt	tgacnttgcc	ttatntnttat	ttntcagaaa	agctntntta	gntctgntgnt	2280
ctntcattatt	tgntntcccat	ctccatnttg	aatntgaatca	ttntgctntcgt	gntcacaata	2340
caatnttagnt	agntacatgc	atntgntcaga	ttcacgntt	atntatgntcat	gactntntaagnt	2400
catgnttagta	cattacctgc	cacgcatgca	ttatntntgnt	tagatntntgat	aggcaaatnt	2460
ggntgntcaac	aatataaaata	taaaatntgt	ttnttatntta	cgaaaataaca	gntgatcaaaa	2520
caaacagntt	tatctnttatt	aacaagntt	tgntntntgnt	tgatgacgnt	ttntntaagnt	2580
tacgctnttcc	ccctntctntt	gaatntntagaa	cactntntatca	tcatntntaatc	aaactntntaa	2640
aaaatntcat	atntntcataaa	taataacaca	aatatntntta	aaaaatctga	aataaataatg	2700
aacaatntta	catntntatca	cgaaaatntca	ttatntntaaaa	tattntntataa	ataaaatgnta	2760
atagnttagntta	tatgnttagntaa	aaaagntactg	cacgcatntaat	atntacaaaa	agatntntaaat	2820
gaactntntat	aaataaataac	actntaaatntaa	tgntntntgnt	tgatgacgnt	ttntntaagnt	2880
aaataaaatnt	tgntaaatntaac	ttctatntatgt	atntntacaca	caaaataaa	ataatntagntaa	2940
aaaaaantntat	gataaantntat	tacctntctca	taagatntntt	aaaataatntga	taaaantntata	3000
gattntntntt	tatgntcaacta	gctagccaaa	aagagaacac	ggnttatntat	aaaaagntga	3060
ccntntntaaat	ctactntgact	tcctntntatc	ctgacgntntt	tatntcaagnt	ggacntacntg	3120
gaagntntntta	atntntcagntc	taaaatntntc	atntntcagct	aaactntnttc	tgntntntatc	3180

-continued

---

ctatcctata agtagtcccg attctcccaa cattgcttat tcacacaact aactaagaaa	3240
gtcttocata gcccccaag cggcccatgg cctcctccga ggacgtcatc aaggagtca	3300
tgcgcttcaa ggtgcgatg gagggctccg tgaacggcca cgagttcgag atcgagggcg	3360
agggcgaggg cgcctctac gagggcacc agaccgcca gctgaaggtg accaagggcg	3420
gccccctgcc cttegcctgg gacatcctgt cccccagtt ccagtacggc tccaaggtgt	3480
acgtgaagca ccccgccgac atccccgact acaagaagct gtccctcccc gagggcttca	3540
agtgggagcg cgtgatgaac ttcgaggacg gcggcgtggt gaccgtgacc caggactcct	3600
ccctgcagga cggctccttc atctacaagg tgaagttcat cggcgtgaac ttcccctccg	3660
acggccccgt aatgcagaag aagactatgg gctgggaggg ctcaccgag cgectgtacc	3720
cccgcgacgg cgtgctgaag ggcgagatcc acaaggcct gaagctgaag gacggcggcc	3780
actacctggt ggagttaag tccatctaca tggccaagaa gcccgtcag ctgccccgct	3840
actactacgt ggactccaag ctggacatca cctcccacia cgaggactac accatcgtgg	3900
agcagtacga gcgcgccgag ggcgccacc acctgttctt gtacggccg gccgcgacac	3960
aagtgtgaga gtactaaata aatgctttgg ttgtacgaaa tcattacact aaataaata	4020
atcaaagctt atatatgctt tccgctaagg ccgaatgcaa agaaattggt tctttctcgt	4080
tatcttttgc cacttttact agtacgtatt aattactact taatcatctt tgtttacggc	4140
tcattatata cgtcgacggc gcgcgcctct agagggccca attcgcccta tagtgagtcg	4200
tattacaatt cactggcogt cgttttacia cgtcgtgact gggaaaaacc tggcgttacc	4260
caacttaatc gccttgacgc acatccccct ttcgccagct ggcgtaatag cgaagaggcc	4320
cgaccgac gcccttccca acagttgcgc agcctatacg tacggcagtt taaggtttac	4380
acctataaaa gagagagcgc ttatcgtctg tttgtggatg tacagagtga tattattgac	4440
acgccggggc gacggatggt gatccccctg gccagtgac gtctgctgtc agataaagtc	4500
tcccgtgaac tttaccgggt ggtgcatac ggggatgaaa gctggcgcac gatgaccacc	4560
gatatggcca gtgtgccogt ctccgttacc ggggaagaag tggctgatct cagccaccgc	4620
gaaaatgaca tcaaaaacgc cattaacctg atgttctggg gaatataat gtcaggcatg	4680
agattataca aaaggatcct cacctagatc cttttcagct agaaagccag tccgcagaaa	4740
cggctctgac cccggatgaa tgtcagctac tgggctatct ggacaaggga aaacgcaagc	4800
gcaaagagaa agcaggtagc ttgcagtggg cttacatggc gatagctaga ctgggcggtt	4860
ttatggacag caagcgaacc ggaattgcca gctggggcgc cctctggtaa ggttgggaag	4920
ccctgcaaag taaactggat ggtttcttg ccgccaagga tctgatggcg caggggatca	4980
agctctgac aagagacagg atgaggatcg tttcgcata ttgaacaaga tggattgac	5040
gcaggttctc cggccgcttg ggtggagagg ctattcggct atgactgggc acaacagaca	5100
atcggctgct ctgatgccg cgtgttccgg ctgtcagcgc aggggcgccc ggttcttttt	5160
gtcaagaccg acctgtccgg tgccctgaat gaactgcaag acgaggcagc gcggctatcg	5220
tggctggcca cgacggcgtt tccttgccga gctgtgctcg acgttgtcac tgaagcggga	5280
agggactggc tgctattggg cgaagtgccg gggcaggatc tcctgtcatc tcacctgct	5340
cctgcccaga aagtatccat catggctgat gcaatgccc ggctgcatac gcttgatccg	5400
gtacctgcc cattcgacca ccaagcgaac catcgcacg agcgagcagc tactcggatg	5460
gaagccggtc ttgtcgatca ggatgatctg gacgaagagc atcaggggct cgcgccagcc	5520
gaactgttcg ccaggctcaa ggcgagcatg cccgacggcg aggatctcgt cgtgacctat	5580

-continued

---

ggcgatgcct gcttgccgaa tatcatgggtg gaaaatggcc gcttttctgg attcatcgac	5640
tgtggccggc tgggtgtggc ggaccgctat caggacatag cgttggtctac cegtgatatt	5700
gctgaagagc ttggcggcga atgggctgac cgcttctctg tgctttaacgg tatcgccgct	5760
cccgatctgc agcgcctcgc cttctatcgc cttcttgaag agttcttctg aattattaac	5820
gcttacaatt tcctgatgcg gtattttctc cttacgcctc tgtgcgggtat ttcacaccgc	5880
atcaggtggc acttttcggg gaaatgtgcg cggaaacctt atttgtttat ttttctaact	5940
acattcaaat atgtatccgc tcatgagaca ataacctga taaatgcttc aataatagca	6000
cgtgaggagg gccaccatgg ccaagttgac cagtgcctt cgggtgctca ccgcgcgca	6060
cgctcccgga gcggtcagct tctggaccga ccggctcggg ttctcccggg acttcctgga	6120
ggacgacttc gccggtgtgg tccgggacga cgtgacctg ttcacagcg cggctccagga	6180
ccagtggtg cgggacaaca ccttgccctg ggtgtgggtg cgcggcctgg acgagctgta	6240
cgccgagtgg tccgaggtcg tgtccacgaa cttccgggac gcctccgggc cggccatgac	6300
cgagatcggc gagcagccgt gggggcggga gttccctctg cgcgacctg cgggcaactg	6360
cgtgcacttc gtggccgagg agcaggactg acacgtgcta aaacttcatt ttaatttaa	6420
aaggatctag gtgaagatcc tttttgataa tctcatgacc aaaatccctt aacgtgagtt	6480
ttcgttccac tgagcgtcag accccgtaga aaagatcaaa ggatcttctt gagatccttt	6540
ttttctgcgc gtaatctgct gcttgcaaac aaaaaacca ccgctaccag cgggtggttg	6600
tttgccgat caagagctac caactctttt tccgaaggtta actggcttca gcagagcgca	6660
gataccaact actgttcttc tagttagacc gtagttaggc caccacttca agaactctgt	6720
agcaccgctt acatacctcg ctctgctaact cctgttacca gtggctgctg ccagtggcga	6780
taagtcgtgt cttaccgggt tggactcaag acgatagtta ccggataagg cgcagcggtc	6840
gggctgaaag gggggtctgt gcacacagcc cagcttgag cgaacgaact acaccgaact	6900
gagataccta cagcgtgagc tatgagaaag cgccacgctt cccgaaggga gaaaggcggga	6960
caggtatccg gtaagcggca gggctcggaa aggagagcgc acgagggagc ttccaggggg	7020
aaacgcctgg tatctttata gtctgtctgg gtttcgccac ctctgacttg agcgtcagtt	7080
tttgtgatgc tcgtcagggg ggcggagcct atggaaaaac gccagcaacg cggccttttt	7140
acggttctcg gccttttctg ggccttttgc tcacatgttc tttctgctg tatcccctga	7200
ttctgtggat aaccgtatta ccgcctttga gtgagctgat accgctcgc gcagccgaac	7260
gaccgagcgc agcagatcag tgagcagga agcggaaagag cgccaatac gcaaacccgc	7320
tctcccgcgc cgttggccga ttcattaatg cagctggcac gacaggttcc ccgactggaa	7380
agcgggcagt gagcgcacg caattaatgt gagttagctc actcattagg caccaccaggc	7440
tttaccttt atgcttccgg ctctgatgtt gtgtggaatt gtgagcggat aacaatttca	7500
cacaggaaac agctatgacc atgattacgc caagctattt aggtgacgcg ttagaatact	7560
caagctatgc atcaagcttg gtaccgagct cggatccact agtaacggcc gccagtgctc	7620
tggaattcag gggcgcgccc tatagatggg atgaagctgc tctcgacaaa tctgataaaa	7680
ctaaagaagg ttagtaatca atttttacaa aatcatagat tttttttttc attgaattat	7740
ttttatgcta taccaagaat tgtattttag tttttgtttt aactacatat aatagaatta	7800
actacatata aattaactaa acttaaaata aaaatagatt tggttcctga aattatttta	7860
agaatatata tgtatatatc taaaatotta gacttagata gatttttcta tctatctatt	7920

-continued

---

ttggttactt	aaaataaata	aatttgtata	aataattgta	tagttatcaa	aaattaanaac	7980
taattttttt	aaagttgttg	atatataaaa	tactaaagat	ttaacgatta	agtatttatt	8040
taagtataga	atthttgttt	ttttttaagt	ttagttatga	agttgttaat	tatattaana	8100
caaaacaata	tttcgaaatt	ttattatcat	attcgaatat	atthttttta	gtgatgatgt	8160
atgaattatt	atcataatth	gaaagtttac	taaaaaatat	atcaacatga	attgtaatat	8220
atgagttatt	accttaacca	aaattataaa	ttaacattaa	atataattat	atatgtcata	8280
tttagccata	caatgtgtca	tcaatattaa	tagtcatgtc	aatattacat	aatgccaata	8340
ttatgctact	taaaccctaa	atcccctaac	tcccgttaa	tagccaaatt	cataaatata	8400
cttattcgac	aaaataaaaa	actttaaaat	atthtactaat	ccgaccatgc	acaagcatcc	8460
attccttatt	ccattgccc	gggataaaaa	tgcaaccnac	tcctcaaaaa	aagaaaaatt	8520
caagctcttt	tgcaaaaaaa	aataaaaaaa	ttttaacacc	taaaatthtt	tgthttccaaa	8580
cttctacagg	gaacacacat	aaaagaaaa	gaggacgtcc	actcggatca	cgcaacaaa	8640
caaaaggtgt	gtcatgactc	ctaagatata	atatttcctt	attcaaaatc	ataccattht	8700
aaattatgaa	tgtatttctg	agtcaccacg	atatgtaatc	caccagcgtt	caaaccaaa	8760
ttttatgatt	gtaagtttaa	gtgaattata	ataatatatt	cttcacggta	tctthttcata	8820
actaattgag	ttatcaaaat	tgatcgaca	tgtggctttg	ataggtgtga	ctthttatggt	8880
atacaattct	ttcaacctaa	aaacattatt	gttcctcaat	atcttacatt	atgcttgact	8940
gcaacaaaat	atthttctcat	ctgtthttctt	cctthtaaac	aatthattat	catctattht	9000
ctgacattht	aatccatcca	cctatgtcaa	aaacttatag	aaaatgtcaa	cttccaaaca	9060
aaacataatt	gaacttcgca	aataaattct	taataatatt	aaaaaatggt	acttaattat	9120
ttcttcaacc	ccatthttcog	cgcgtagcgc	ggacaaagac	tctagttaaa	tatagaagtt	9180
tccgattctc	atcgtataaa	acggtgactt	tggcgggctt	tcatgtgtaa	caaattggtt	9240
taacaaaacca	ctgcctagtc	gtthtagtga	gaatcagcgc	atggaaactc	gattggagcg	9300
tgactthtca	gtgcgggagg	cccaccacca	cagcggggct	tacgtctcaa	gaatctcgcc	9360
cacggtthtc	ttcatctgcc	ccccgccaag	tgtcttctct	gttcggcaact	tctcaccag	9420
ttacaggaac	cctaaaaaatg	gcctthtctc	agccccggct	ataatacaca	catgatccta	9480
tagtgggttc	ttccacaagt	tacatctcct	tctggattgt	acatthtcaag	tgtthtgtgt	9540
ttthtctgct	ctgagagaaa	atcgcggccg	catggagaga	tctcaacggc	agtctcctcc	9600
gccaccgtcg	ccgtctctct	cctcgtctct	cgtctcccg	gacaccgtcc	tcgtccctcc	9660
cggaaaagagg	cggagggcgg	cgacggccaa	ggccggcgcc	gagcctaata	agaggatccg	9720
caaggacccc	gccgccgccg	ccgcggggaa	gaggagctcc	gtctacaggg	gagtcaccag	9780
gcacaggtgg	acgggcaggt	tcgaggcgca	tctctgggac	aagcaactgcc	tcgcccgct	9840
ccacaacaag	aagaaaggca	ggcaagtcta	cctggggggc	tatgacagcg	aggaggcagc	9900
tgctcgtgcc	tatgacctcg	cagctctcaa	gtactggggt	cctgagactc	tgctcaactt	9960
ccctgtggag	gattactcca	gcgagatgcc	ggagatggag	gccgtgtccc	gggaggagta	10020
cctggcctcc	ctccgccgca	ggagcagcgg	cttctccagg	ggcgtctcca	agtagagagg	10080
cgtegccagg	catcaccaca	acgggaggtg	ggaggcacgg	attgggcgag	tctthtggaa	10140
caagtacctc	tacttgggaa	catttgacac	tcaagaagag	gcagccaagg	cctatgacct	10200
tgccgccatt	gaataccgtg	gcgtcaatgc	tgtaaccaac	ttcgacatca	gctgctacct	10260
ggaccaccog	ctgttctctg	cacagctcca	acaggagcca	caggtggtgc	cggcactcaa	10320

-continued

---

```

ccaagaacct caacctgatc agagcgaaac cggaactaca gagcaagagc cggagtcaag 10380
cgaagccaag acaccggatg gcagtgacaga acccgatgag aacgcgggtgc ctgacgacac 10440
cgcgagagccc ctcaccacag tgcagcagac catcgaagag ggcttgtgga gcccttgcac 10500
ggattacgag ctagacacca tgtcgagacc aaactttggc agctcaatca atctgagcga 10560
gtggttcgct gacgcagact tgcactgcaa catcggatgc ctgttcgatg ggtgttctgc 10620
ggctgacgaa ggaagcaagg atgggtgtagg tctggcagat ttcagtctgt ttgaggcagc 10680
tgatgtccag ctgaaggatg ttctttcgga tatggaagag gggatacaac ctccagcgat 10740
gatcagtgtg tgcaacgc                                     10758

```

```

<210> SEQ ID NO 71
<211> LENGTH: 20553
<212> TYPE: DNA
<213> ORGANISM: Artificial sequence
<220> FEATURE:
<223> OTHER INFORMATION: Vector ARAL079
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: (14134)..(14134)
<223> OTHER INFORMATION: n is a, c, g, or t
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: (18743)..(18743)
<223> OTHER INFORMATION: n is a, c, g, or t

<400> SEQUENCE: 71

```

```

cgcgctcga gtgggcgat cccccggct gcaggaattc actggccgtc gttttacaac 60
gtcgtgactg ggaaaacct ggcggtacc aacttaatcg ccttgacga catccccctt 120
tcgccagctg gcgtaatagc gaagaggccc gcaccgatcg cccttccaa cagttgagca 180
gcctgaatgg cgaatggatc gatccatgc gatgtacctt ttgttagtca gcctctcgat 240
tgctcatcgt cattacacag taccgaagt tgatcgatct agtaacatag atgacaccgc 300
gcgcgataat ttatcctagt ttgcgcgcta tttttgttt tctatcgcgt attaatgta 360
taattgctgg actctaata taaaaacca tctcataaat aacgtcatgc attacatgtt 420
aattattaca tgettaacgt aattcaacag aaattatatg ataatcatcg caagaccggc 480
aacaggattc aatcttaaga aactttattg ccaaatgttt gaacgatctg cttcgaagca 540
ctccttcttt actccaccat ctgcctctta ttgaaaacgt gggtagcacc aaaacgaatc 600
aagtcgctgg aactgaagt accaatcacg ctggatgatt tgccagttgg attaactctg 660
cctttccccg catgaataat attgatgaat gcatgcgtga gggtagttc gatggtggca 720
atagctgcaa ttgccgcgac atcctccaac gagcataatt cttcagaaaa atagcgatgt 780
tccatgttgt cagggcatgc atgatgcacg ttatgaggtg acggtgctag gcagatttcc 840
ctcaaagttt catagtcagt atcatattca tcattgcatt cctgcaagag agaattgaga 900
cgcaatccac acgctgctgc aaccttcggt cgttcgtggt ctatttgcct ttggacgttg 960
caaacgtaag tgttggatcg atccgggtg ggccaagaac tccagcatga gatccccgag 1020
ctggaggatc atccagcggc cgtcccgtaa aacgattccg aagoccaacc tttcatagaa 1080
ggcggcggtg gaatcgaat ctctgatgag caggttgggc gtcgcttggc cggtcatttc 1140
gaaccccaga gtcccgtca gaagaactcg tcaagaaggc gatagaaggc gatgcgctgc 1200
gaatcgggag cggcgatacc gtaaacgacg aggaagcggc cagccattc gccgccaagc 1260
tcttcagcaa taccacgggt agccaacgct atgtctgat agcggctcgc cacaccagc 1320

```

-continued

---

cggccacagt	cgatgaatcc	agaaaagcgg	ccattttcca	ccatgatatt	cggaagcag	1380
gcatcgccat	gggtcaegac	gagatcctcg	ccgtcgggca	tgcgcgcctt	gagcctggcg	1440
aacagttegg	ctggcgcgag	cccctgatgc	tcttctgcca	gatcatcctg	atcgacaaga	1500
ccggcttcca	tccgagtagc	tgctcgcctg	atgcatggtt	tgccttggtg	gtcgaatggg	1560
caggtagccg	gatcaagcgt	atgcagccgc	cgcattgcat	cagccatgat	ggatactttc	1620
tggcaggag	caaggtgaga	tgacaggaga	tctgccccg	gcacttcgcc	caatagcagc	1680
cagtcccttc	ccgcttcagt	gacaacgtcg	agcacagctg	cgcaaggaac	gcccgctcgtg	1740
gccagccacg	atagccgcgc	tgctcgtcc	tgcatctcat	tcagggcacc	ggacaggctc	1800
gtcttgacaa	aaagaaccgg	gccccctgc	gctgacagcc	ggaacacggc	ggcatcagag	1860
cagccgattg	tctgttgctg	ccagtcatag	ccgaatagcc	tctccacca	agcggccgga	1920
gaacctgcgt	gcaatccatc	ttgttcaatc	atgcgaaacg	atccccgcaa	gcttgagagc	1980
tggtgatttc	agcgtgtcct	ctccaaatga	aatgaacttc	cttatataga	ggaagggctc	2040
tgcaaggat	agtgggattg	tgctcctcc	cttacgtcag	tggagatata	acatcaatcc	2100
acttgctttg	aagacgtgg	tggaaactct	tctttttcca	cgatgctcct	cgtgggtggg	2160
ggctccatct	tgggaccact	gtcggcagag	gcattctcaa	cgatggcctt	tcctttatcg	2220
caatgatggc	attttagtag	gccaccttcc	ttttcacta	tcttcacaat	aaagtgcagc	2280
atagctgggc	aatggaatcc	gaggagggtt	ccggatatta	ccctttgttg	aaaagtctca	2340
attgcccttt	ggctctctga	gactgtatct	ttgatatttt	tggagtagac	aagcgtgctg	2400
tgctccacca	tgttgacgaa	gattttcttc	ttgtcattga	gtcgtaaagag	actctgtatg	2460
aactgttcgc	cagctcttac	ggcagattct	gttaggtcct	ctatttgaat	ctttgactcc	2520
atggcctttg	attcagtggt	aactacctt	ttagagactc	caatctctat	tacttgctct	2580
ggtttgtaa	gcaagccttg	aatcgtccat	actggaatag	tacttctgat	cttgagaaat	2640
atatctttct	ctgtgtctct	gatgcagtta	gtcctgaatc	ttttgactgc	atctttaacc	2700
ttcttgggaa	ggatattgat	ctcctggaga	ttattgctcg	ggtagatcgt	cttgatgaga	2760
cctgctgcgt	aagcctctct	aacctctgt	gggttagcat	tctttctgaa	attgaaaagg	2820
ctaactctct	cattatcagt	ggatgaacatg	gtatcgtcac	cttctccgct	gaacttccctg	2880
actagatcgt	agagatagag	gaagtcgtcc	attgtgatct	ctggggcaaa	ggagatctga	2940
attaatcga	tatggtggat	ttatcacaaa	tgggaccgcg	cgccgacaga	ggtgtgatgt	3000
taggccagga	ctttgaaat	ttgcgcaact	atcgtatagt	ggccgacaaa	ttgacccgca	3060
gttgacagac	tgcttagcat	ttgagtgaat	tatgtgaggt	aatgggctac	actgaattgg	3120
tagctcaaac	tgctcagatt	tatgtatatg	agtgtatatt	ttcgcataat	ctcagaccaa	3180
tctgaagatg	aaatgggtat	ctgggaatgg	cgaaatcaag	gcatcgcctg	tgaagtttct	3240
catctaagcc	cccatttggg	cgtgaatgta	gacacgtcga	aataaagatt	tccgaattag	3300
aataatttgt	ttattgcttt	cgctataaaa	tacgacggat	cgtaatttgt	cgttttatca	3360
aaatgtactt	tcattttata	ataacgctgc	ggacatctac	atttttgaat	tgaaaaaaaa	3420
ttggtaatta	ctctttcttt	ttctccatat	tgaccatcat	actcattgct	gatccatgta	3480
gatttcccgg	acatgaagcc	atttacaatt	gaatatatcc	tgcgcccgtc	gcccgtttgc	3540
accgggtgga	gcttgcatgt	tggtttctac	gcagaactga	gccgggttagg	cagataattt	3600
ccattgagaa	ctgagccatg	tgacacttcc	ccccaacacg	gtgagcgacg	gggcaacgga	3660
gtgatccaca	tgggactttt	aaacatcatc	cgtcggatgg	cgttgcgaga	gaagcagctg	3720

-continued

---

atccgtgaga	tcagccgacg	caccgggacg	gcgcgcaaca	cgatcgcaaa	gtatttgaac	3780
gcaggtacaa	tcgagccgac	gttcacgcgg	aacgaccaag	caagctagct	ttaatgcggt	3840
agtttatcac	agttaaattg	ctaaccgagt	caggcaccgt	gtatgaaatc	taacaatgcg	3900
ctcatcgtea	tcctcggcac	cgtcacccctg	gatgctgtag	gcataggctt	ggttatgccc	3960
gtactgcccg	gcctcttgcg	ggatatcgtc	cattccgaca	gcatcgccag	tcactatggc	4020
gtgctgctag	cgctatatgc	gttgatgcaa	tttctatgcg	caccocgttct	cggagcactg	4080
tccgaccgct	ttggccgccc	cccagtcctg	ctcgettcgc	tacttgagagc	cactatcgac	4140
tacgcgatca	tggcgaccac	accocgtcctg	tggtccaacc	cctccgctgc	tatagtgcag	4200
tcggcttctg	acgttcagtg	cagccgtctt	ctgaaaacga	catgtcgcac	aagtccctaag	4260
ttacgcgaca	ggctgcgcc	ctgccctttt	cctggcgttt	tcttgctcgcg	tgttttagtc	4320
gcataaagta	gaataactgc	gactagaacc	ggagacatta	cgccatgaac	aagagcgccg	4380
ccgctggcct	gctgggtat	gcccgcgca	gcaccgacga	ccaggacttg	accaaccaac	4440
gggcccgaact	gcaecgccc	ggctgcacca	agctgttttc	cgagaagatc	accggcacca	4500
ggcgcgacgc	cccggagctg	gccaggatgc	ttgaccacct	acgcocctggc	gacgttgga	4560
cagtgaccag	gctagaccgc	ctggcccgc	gcaccocga	cctactggac	attgcccagc	4620
gcatccagga	ggccggcgcg	ggcctgcgta	gacctggcaga	gacctggccc	gacaccacca	4680
cgccggcccg	ccgcatggtg	ttgaccgtgt	tcgccggcat	tgccgagttc	gagcgttccc	4740
taatcatcga	ccgcaccocg	agcgggcccg	aggcccacaa	ggcccagggc	gtgaagtttg	4800
gccccgccc	tacctcacc	ccggcacaga	tcgcgcacgc	ccgcgagctg	atcgaccagg	4860
aaggcccgc	cgtgaaagag	gcggctgcac	tgcttggcgt	gcatcgctcg	accctgtacc	4920
gcgcacttga	gcgcagcgag	gaagtgcgc	ccaccgaggc	caggcggcgc	ggtgccttc	4980
gtgaggacgc	attgaccgag	gcccagccc	tggcggccc	cgagaatgaa	cgccaagagg	5040
aacaagcatg	aaaccgcacc	aggacggcca	ggacgaaccg	ttttcatta	ccgaagagat	5100
cgaggcggag	atgatcgcgg	ccgggtacgt	gttcgagccg	cccgcgcacg	tctcaaccgt	5160
gcggctgcat	gaaatcctgg	ccggtttgtc	tgatgccaag	ctggcggcct	ggcggccag	5220
cttgccgct	gaagaaaccg	agcgcgcgcg	tctaaaagg	tgatgtgtat	ttgagtaaaa	5280
cagcttgcgt	catcggctcg	ctcgtatat	gatgcgatga	gtaaataaac	aaatacgcaa	5340
gggaacgcat	gaagtatcgc	ctgtacttaa	ccagaaaggc	gggtcaggca	agacgaccat	5400
cgcaaccat	ctagcccgcg	ccctgcaact	cgccggggcc	gatgttctgt	tagtcgattc	5460
cgatccccag	ggcagtgccc	gcgattgggc	ggcctgccc	gaagatcaac	cgctaaccgt	5520
tgccggcatc	gaccgcccga	cgattgaccg	cgacgtgaag	gccatcggcc	ggcgcgactt	5580
cgtagtgatc	gacggagcgc	cccaggcggc	ggacttgct	gtgtccgcga	tcaaggcagc	5640
cgacttcgtg	ctgattccgg	tcagccaag	cccttacgac	atatgggcca	ccgccgacct	5700
ggtggagctg	gttaagcagc	gcattgaggt	cacggatgga	aggctacaag	cgccctttgt	5760
cgtgtcgcgg	gcgatcaaac	gcacgcgcac	cgccgggtgag	gttgccgagg	cgctggcccg	5820
gtacgagctg	cccattcttg	agtcctgat	cacgcagcgc	gtgagctacc	caggcactgc	5880
cgcccccgcg	acaaccgttc	ttgaatcaga	accggagggc	gacgctgccc	gcgaggtcca	5940
ggcgtggcc	gctgaaatta	aatcaaaact	catttgagtt	aatgaggtaa	agagaaaatg	6000
agcaaaagca	caaacacgct	aagtgcggc	cgtccgagcg	cacgcagcag	caaggetgca	6060

-continued

---

acgttgcca gcctggcaga cacgccagcc atgaagcggg tcaactttca gttgccggcg	6120
gaggatcaca ccaagctgaa gatgtacgcg gtacgccaaag gcaagaccat taccgagctg	6180
ctatctgaat acatcgcgca gctaccagag taaatgagca aatgaataaa tgagtagatg	6240
aattttagcg gctaaaggag gcggcatgga aaatcaagaa caaccaggca ccgacgccgt	6300
ggaatgcccc atgtgtggag gaacggggcg ttggccaggc gtaagcggct gggttgtctg	6360
ccggccctgc aatggcactg gaacccccaa gcccgaggaa tcggcgtgag cggtcgcaaa	6420
ccatccggcc cgggtacaaat cggcgcggcg ctgggtgatg acctggtgga gaagttgaag	6480
gccgcgcagg ccgcccagcg gcaacgcatac gaggcagaag cacgcccgg tgaatcgtgg	6540
caagcggccg ctgategaat ccgcaaagaa tcccggcaac cgccggcagc cggtgccgccg	6600
tcgattagga agccgccccaa gggcgacgag caaccagatt ttttcgttcc gatgctctat	6660
gacgtgggca ccccgcatag tcgcagcatac atggacgtgg ccgttttccg tctgtcgaag	6720
cgtgaccgac gagctggcga ggtgatccgc tacgagcttc cagacgggca cgtagaggtt	6780
tccgcagggc cggccggcat gggcagtggt tgggattacg acctggtact gatggcgggt	6840
tcccatctaa ccgaatccat gaaccgatac cgggaaggga agggagacaa gcccggccgc	6900
gtgttccgtc cacacgttgc ggagctactc aagttctgcc ggcgagccga tggcggaaaag	6960
cagaaagacg acctggtaga aacctgcatt cggttaaaca ccacgcacgt tgccatgcag	7020
cgtacgaaga aggccaaagaa cggccgcctg gtgacgggat ccgaggggtga agccttgatt	7080
agccgctaca agatcgtaaa gagcgaacc gggcggccgg agtacatcga gatcgagcta	7140
gctgattgga tgtaccgca gatcacagaa ggcaagaacc cggacgtgct gacggttcac	7200
cccgattact ttttgatcga tcccggcatac ggccgttttc tctaccgctt gccacgccgc	7260
gccgcaggca aggcagaagc cagatgggtg ttcaagacga tctacgaacg cagtggcagc	7320
gccggagagt tcaagaagtt ctgtttcacc gtgcgcaagc tgatcgggtc aatgacctg	7380
ccggagtacg atttgaagga ggaggcgggg caggctggcc cgatcctagt catgcgctac	7440
cgcaacctga tcgagggcga agcatccgcc ggttccaat gtacggagca gatgctaggg	7500
caaattgccc tagcagggga aaaaggtcga aaaggtctct ttctgtgga tagcagctac	7560
attgggaacc caaagccgta cattgggaac cggaaacctg acattgggaa cccaaagccg	7620
tacattggga accggtcaca catgtaagtg actgatataa aagagaaaaa aggcgatttt	7680
tccgcctaaa actctttaa acttattaaa actcttaaaa cccgcctggc ctgtgcataa	7740
ctgtctggcc agcgcacagc cgaagagctg caaaaagcgc ctacccttcg gtcgctgcgc	7800
tccctacgcc ccgccccttc gcgtcggcct atcgcggccg ctggccgctc aaaaatggct	7860
ggcctacggc caggcaatct accagggcgc ggacaagccg ccgcgctgcc actcgaccgc	7920
cggcgcacc accaaggcac cctgcctcgc cggtttcggt gatgacgggtg aaaaacctctg	7980
acacatgcag ctcccggaga cggtcacagc ttgtctgtaa gcggatgccg ggagcagaca	8040
agcccgtcag ggcgcgtcag cgggtgttgg cgggtgtcgg ggcgcagcca tgaccagtc	8100
acgtagcgat agcggagtgt atactggctt aactatgcgg catcagagca gattgtactg	8160
agagtgcacc atatgcgggtg tgaatatccg cacagatgcg taaggagaaa ataccgcatc	8220
aggcgtcttt ccgcttctc gctcactgac tcgctgcgct cggtcgttcg gctgcggcga	8280
gcggtatcag ctcaactcaa ggcggtaata cggttatcca cagaatcagg ggataacgca	8340
ggaaagaaca tgtgagcaaa aggccagcaa aaggccagga accgtaaaaa ggccgcggtg	8400
ctggcgtttt tccataggct ccgccccct gacgagcacc acaaaaatcg acgctcaagt	8460

-continued

---

cagaggtggc	gaaacccgac	aggactataa	agataaccagg	cgtttcccc	tggaagctcc	8520
ctcgtgctct	ctcctgttcc	gacctgccc	cttaccggat	acctgtccc	ctttctccct	8580
tcgggaagcg	tggcgcttcc	tcatagctca	cgctgtaggt	atctcagttc	ggtgtaggtc	8640
gttcgctcca	agctgggctg	tgtgcacgaa	cccccgctc	agcccgaccg	ctgcccctta	8700
tccgtaact	atcgtcttga	gtccaacccg	gtaagacacg	acttatcgcc	actggcagca	8760
gccactggta	acaggattag	cagagcgagg	tatgtaggcg	gtgctacaga	gttcttgaag	8820
tggtggccta	actacggcta	cactagaagg	acagtatttg	gtatctgcgc	tctgctgaag	8880
ccagttacct	tcggaaaaag	agttggtagc	tcttgatccg	gcaaaaaaac	caccgctggt	8940
agcgtgggtt	tttttggttg	caagcagcag	attacgcgca	gaaaaaaagg	atctcaagaa	9000
gatcctttga	tcttttctac	ggggtctgac	gctcagtgga	acgaaaactc	acgttaaggg	9060
atthttggtca	tgagattatc	aaaaaggatc	ttcacctaga	tccttttaaa	ttaaaaaatga	9120
agtttttaaat	caatctaag	tatatatgag	taaacttggt	ctgacagtta	ccaatgctta	9180
atcagtgagg	cacctatctc	agcgatctgt	ctatttctgt	catccatagt	tgctgactc	9240
cccgtcgtgt	agataactac	gatacgggag	ggcttaccat	ctggcccag	tgctgcaatg	9300
ataccgcgag	accacgcctc	accggctcca	gatttatcag	caataaacca	gccagccgga	9360
agggccgagc	gcagaagtgg	tctgcaact	ttatccgctc	ccatccagtc	tattaattgt	9420
tgccgggaag	ctagagtaag	tagttcgcca	gttaatagtt	tgcccaactg	tgttgccatt	9480
gctacaggca	tcgtggtgtc	acgctcgtcg	tttggtatgg	cttcattcag	ctccggttcc	9540
caacgatcaa	ggcgagttac	atgatcccc	atggttgca	aaaaagcggg	tagctccttc	9600
ggtcctccga	tcgttgcag	aagtaagttg	gccgcagtg	tatcactcat	ggttatggca	9660
gcactgcata	attctcttac	tgctatgcca	tccgtaagat	gcttttctgt	gactggtgag	9720
tactcaacca	agtcattctg	agaatagtg	atgcccgcag	cgagttgctc	ttgcccggcg	9780
tcaacacggg	ataataccgc	gccacatagc	agaactttaa	aagtgctcat	cattggaaaa	9840
gacctgcagg	gggggggggg	cgctgaggtc	tgccctcgta	agaaggtggt	gctgactcat	9900
accaggcctg	aatcgcccc	tcaccagcc	agaaagtgag	ggagccacgg	ttgatgagag	9960
ctttgttgta	ggtggaccag	ttggtgattt	tgaacttttg	ctttgccacg	gaacggtctg	10020
cgttgtcggg	aagatgcgtg	atctgatcct	tcaactcagc	aaaagttcga	tttattcaac	10080
aaagccgccc	tcccgtcaag	tcagcgtaat	gctctgccag	tgttacaacc	aattaaccaa	10140
ttctgattag	aaaaactcat	cgagcatcaa	atgaaactgc	aatttattca	tatcaggatt	10200
atcaatacca	tatttttgaa	aaagccggtt	ctgtaatgaa	ggagaaaact	caccgaggca	10260
gttccatagg	atggcaagat	cctggtatcg	gtctgcgatt	ccgactcgtc	caacatcaat	10320
acaacctatt	aatttcccct	cgtaaaaaat	aaggttatca	agtgagaaat	caccatgagt	10380
gacgactgaa	tccggtgaga	atggcaaaaag	cttatgcatt	tctttccaga	cttggttcaac	10440
agggcagcca	ttacgctcgt	catcaaaaatc	actcgcatac	accaaacctg	tattcattcg	10500
tgattgcgcc	tgagcgagac	gaaataccgc	atcgcgtgta	aaaggacaat	tacaaaacagg	10560
aatcgaatgc	aaccggcgca	ggaacactgc	cagcgcatac	acaatatttt	cacctgaatc	10620
aggatattct	tctaatacct	ggaatgctgt	tttcccgggg	atcgcagtg	tgagtaacca	10680
tgcatcatca	ggagtacgga	taaaatgctt	gatggtcgga	agaggcataa	attccgctcag	10740
ccagtttagt	ctgaccatct	catctgtaac	atcattggca	acgctacctt	tgccatggtt	10800

-continued

---

cagaaacaac	tctggcgc	at	cg	ggcttccc	atacaatcga	tagattgtcg	cacctgattg	10860
cccgacatta	tcgcgagccc	atttataccc	atataaatca	gcatccatgt	tggaatttaa			10920
tcgcggcctc	gagcaagacg	tttcccgttg	aatatggctc	ataacacccc	ttgtattact			10980
gtttatgtaa	gcagacagtt	ttattgttca	tgatgatata	ttttatctt	gtgcaatgta			11040
acatcagaga	ttttgagaca	caacgtggct	ttccccccc	cccctgcagg	tcaattcggt			11100
cgatatggct	attacgaaga	aggctcgtgc	gcgagatccc	gtgaactttc	ccacgcaaca			11160
agtgaaccgc	accgggtttg	ccggaggcca	tttcgttaaa	atgctgcagcc	atggctgctt			11220
cgtccagcat	ggcgaatac	tgatcctcgt	cttcggctgg	cggtatatg	ccgatgggct			11280
tcaaaagccg	ccgtggttga	accagtcctat	ccattccaag	gtagcgaact	cgaccgcttc			11340
gaagctctc	catggctccac	gccgatgaat	gacctcgcc	ttgtaaagac	cgttgatcgc			11400
ttctgcgagg	gctgtgtcgt	gctgtgcgcg	acgcttccga	tagatggctc	gatacctgct			11460
tctgccaaacc	gctcgggaata	gcgaaaggac	acgtattgaa	caccgcgac	cgagtgatgc			11520
actaggccgc	catgagcggg	acgccgatca	tgatgagcct	cctcgagggc	atcgaggaca			11580
aagcctgcat	gtgctgtccg	gctcgcgcc	catccgacaa	tgcgacgggc	gaagacgtcg			11640
atcacgaagg	ccacgtagac	gaagccctcc	caagtggcga	cataagtacg	gacatgcgca			11700
aaggctttcc	cggtttgtcg	ctgatggctc	aagagacgct	gaagcgcgat	ccgatgcgca			11760
ggcatctgtt	cgtcttcgcg	ggctgtggcg	gtggcctgat	caaggcact	cgccgaagag			11820
ctgcatgatt	ggctcgaaac	cgagcggggg	aaattgtcgc	gcagttctcc	cgtcgcgag			11880
gcgataaatt	acatgctcaa	gcgatgggat	ggcattacgt	cattcctcga	tgacggcccg			11940
atttgcctga	cgaacaatgc	tgccgaacga	acgctcagag	gctatgtact	cggcaggaag			12000
tcatggctgt	ttgccggatc	ggatcgttgt	gctgaacgtg	ggcggttcat	ggcgacactg			12060
atcatgagcg	ccaagctcaa	taacatcgat	ccgcaggcct	ggcttgccga	cgtccgcgcc			12120
gaccttgccg	acgctccgat	cagcaggctt	gagcaacagc	tgccgtggaa	ctggacatcc			12180
aagacactga	gtgctcaggc	ggcctgacct	gcgcccttca	ccggatactt	accccattat			12240
cgcagattgc	gatgaagcat	cagcgtcatt	cagcaatctt	gccaaagtat	gcaggctcgc			12300
gagaatcgac	gtgcgaaacc	ggctggttgc	gccaaagatc	cgcttgccga	gcggctgaa			12360
atcatgctg	ggacttcaag	aggctcagta	gaggaagaac	cggaaagggt	gcaccggaaa			12420
atatgcgctc	ctttggagag	cgctcctatg	acgtgaacaa	atcgcccgga	ccaaggatgc			12480
cacggataca	aaagctcgcg	aagctcggtc	ccgtgggtgt	tctgtcgtct	cgttgtataa			12540
cgaaatccat	tcccattccg	cgtcgaagat	ggcttcccct	cggcagttca	tcagggctaa			12600
atcaatctag	ccgacttgtc	cggtgaaatg	ggctgcactc	caacagaaac	aatcaaaaa			12660
acatacacag	cgacttattc	acacgagctc	aaattacaac	ggatatatc	ctgccagtca			12720
gcatcatcac	acccaaaagt	aggcccgaat	agtttgaat	tagaaagctc	gcaattgagg			12780
tctacaggcc	aaattcgtc	ttagccgtac	aatattactc	accggtgcga	tgcccccat			12840
cgtaggtgaa	ggtggaatt	aatgatccat	cttgagacca	caggcccaca	acagctacca			12900
gttccctcaa	gggtccacca	aaaacgtaag	cgcttacgta	catggctgat	aagaaaaggc			12960
aatgtgtaga	tgtaaacatc	caacgtcgc	ttcagggatc	gatccaatac	gcaaacccgc			13020
tctccccgcg	cgttgccga	ttcattaatg	cagctggcac	gacaggttcc	ccgactggaa			13080
agcgggcagt	gagcgaacg	caattaatgt	gagttagctc	actcattagg	caccccaggc			13140
tttacacttt	atgcttccg	ctcgtatgtt	gtgtggaatt	gtgagcggat	aacaatttca			13200

-continued

---

cacaggaaac agctatgacc atgattacgc caagcttgca tgectgcagg tegactctag 13260  
aggatctggc gcgccctata gatgggatga agctgctctc gacaaatctg ataaaaactaa 13320  
agaagggttag taatcaattt ttacaaaatc atagattatt tttttcattg aattatTTTT 13380  
atgctataacc aagaattgta ttttagtatt tgttttaact acatataata gaattaacta 13440  
catataaatt aactaaactt aaaataaaaa tagatttgtt tcctgaaatt attttaagaa 13500  
tatatatgta tatatctaaa atcttagact tagatagatt tttctatcta tctatTTTTg 13560  
ttacttaaaa taaataaatt tgtataaata attgtatagt tatcaaaaaa taaaactaat 13620  
ttttttaag ttgttgat ataaaaact aaagatttaa cgattaagta tttatttaag 13680  
tatagaattt tgtTTTTttt ttaagtttag ttatgaagtt gttaattata ttaaaacaaa 13740  
acaatatttc gaaattttat tatcatatc gaatatatt ttttagtga tgatgtatga 13800  
attattatca taatttgaaa gtttactaaa aaatatatca acatgaattg taatatatga 13860  
gttattacct taacaaaat tataaattaa cattaatat aattatatat gtcataattt 13920  
gccatacaat gtgtcatcaa tattaatagt catgtcaata ttacataatg ccaatattat 13980  
gctacttaaa ccccaaatcc cctaactccc gttaagtagc caaatcata aatatactta 14040  
ttcgacaaaa taaaaaactt taaaatattt actaatccga ccatgcacaa gcatccattc 14100  
cctattccat tgcccaggga taacaatgca accnactcct caaaaaaga aaaattcaag 14160  
ctcttttgca aaaaaaata aaataattt aacacctaaa atttttgtt tccaaacttc 14220  
tacagggaac acacataaaa gaaaaagagg acgtccactc ggatcacgca acaaaccaaa 14280  
agggtgtgca tgactcctaa gatataatat ttccttattc aaaatcatac cattttaaat 14340  
tatgaatgta tttcgtagtc caccagatat gtaatccacc agcgttcaaa ccaaagttt 14400  
atgattgtaa gtttaagtga attataataa tatattcttc acggtatcct ttcataacta 14460  
attgagttat caaacttgat cgcacatgtg gctttgatag gtgtgacttt tatggtatac 14520  
aattctttca acctaaaaac attattgttc ctcaatatct tacattatgc ttgactgcaa 14580  
caaaatattt tctcatctgt tttctctctt taaaccaatt tattatcctc ttttctctga 14640  
cattttaatc catccaccta tgtcaaaaac ttatagaaaa tgtcaacttc caaacaaaac 14700  
ataattgaac ttcgcaaata aattcttaat aatattaaaa atgttactt aattatttct 14760  
tcaaccccat tttccgcgcg tagcgcggac aaagactceta gttaaatata gaagtttccg 14820  
attctcatcg tataaaacgg tgactttggc gggctttcat gtgtaacaaa ttggtttaac 14880  
aaaccactgc ctagtcgttt agtgtagaat cagcgcattg aactccgatt ggagcgtgac 14940  
tttcactgac cggaggccca ccaccacagc gggcgttacg ctctaagaat ctgcgccacg 15000  
gtttctttca tctgcccccc gccaaagtgc ttcctcgttc gccacttctc accaagttac 15060  
aggaacccta aaaatggcct ttcttcagcc ccggctataa tacacacatg atcctatagt 15120  
gggttcttcc acaagttaca tctcctctcg gattgtacat ttcaagtgtt tgtgtttttt 15180  
ctgcctctga gagaaaatcg cggccgcgatg gagagatctc aacggcagtc tctccgccca 15240  
ccgtcgcgct cctcctctc gtcctccgtc tccgaggaca ccgtcctcgt cctcctccga 15300  
aagaggcgga gggcggcgac ggccaaggcc ggcgcccagc ctaataagag gatccgcaag 15360  
gaccccgccg ccgcccgcgc ggggaagagg agctccgtct acaggggagt caccaggcac 15420  
agggtggacgg gcaggttcga ggcgcacttc tgggacaagc actgcctcgc cgcgctccac 15480  
aacaagaaga aaggcaggca agtctacctg gggcggtatg acagcgagga ggcagctgct 15540

-continued

---

cgtgcctatg	acctcgcagc	tctcaagtac	tggggtcctg	agactctgct	caacttcctt	15600
gtggaggatt	actccagcga	gatgccggag	atggaggccg	tgtcccggga	ggagtacctg	15660
gcctccctcc	gccgcaggag	cagcggcttc	tccaggggcg	tctccaagta	cagaggcgtc	15720
gccaggcatc	accacaacgg	gaggtgggag	gcacggattg	ggcgagtctt	tgggaacaag	15780
tacctctact	tgggaacatt	tgacactcaa	gaagaggcag	ccaaggccta	tgaccttgcg	15840
gccattgaat	accgtggcgt	caatgctgta	accaacttcg	acatcagctg	ctacctggac	15900
caccgcgtgt	tcctggcaca	gctccaacag	gagccacagg	tgggtcccgc	actcaaccaa	15960
gaacctcaac	ctgatcagag	cgaaacggga	actacagagc	aagagccgga	gtcaagcgaa	16020
gccaaagacac	cggatggcag	tcgagaaccc	gatgagaacg	cggtgcctga	cgacaccgcg	16080
gagcccctca	ccacagtoga	cgacagcatc	gaagagggct	tgtggagccc	ttgcatggat	16140
tacgagctag	acaccatgtc	gagaccaaac	tttggcagct	caatcaatct	gagcagtggt	16200
ttcgtgacg	cagacttoga	ctgcaacatc	ggatgctgt	tcgatgggtg	ttctgcgget	16260
gacgaaggaa	gcaaggatgg	tgtaggtctg	gcagatttca	gtctgtttga	ggcaggtgat	16320
gtccagctga	aggatgttct	ttcggatatg	gaagagggga	tacaacctcc	agcagatgatc	16380
agtgtgtgca	acgcggccgc	aagatgaac	taaaatgcat	gtaggtgtaa	gagctcatgg	16440
agagcatgga	atattgtatc	cgaccatgta	acagtataat	aactgagctc	catctcactt	16500
cttctatgaa	taaacaagg	atgttatgat	atattaacac	tctatctatg	caccttattg	16560
ttctatgata	aatttctctt	tattattata	aatcatctga	atcgtgacgg	cttatggaat	16620
gcttcaaata	gtacaaaaac	aaatgtgtac	tataagactt	tctaaacaat	tctaacctta	16680
gcattgtgaa	cgagacataa	gtgttaagaa	gacataacaa	ttataatgga	agaagtttgt	16740
ctccatttat	atattatata	ttaccactt	atgtattata	ttaggatggt	aaggagacat	16800
aacaattata	aagagagaag	tttgtatcca	tttatatatt	atatactacc	catttatata	16860
ttatacttat	ccacttattt	aatgtcttta	taaggtttga	tccatgatat	ttctaataat	16920
ttagttgata	tgtatatgaa	agggtactat	ttgaactctc	ttactctgta	taaaggttgg	16980
atcatcctta	aagtgggtct	atttaatttt	attgcttctt	acagataaaa	aaaaaattat	17040
gagttggttt	gataaaatat	tgaaggattt	aaaataataa	taaataacat	ataatatatg	17100
tatataaatt	tattataata	taacatttat	ctataaaaaa	gtaaatattg	tcataaatct	17160
atacaatcgt	ttagccttgc	tggacgaatc	tcaattattt	aaacgagagt	aaacatattt	17220
gactttttgg	ttatttcaac	aattattatt	taacactata	tgaaattttt	ttttttatca	17280
gcaaagaata	aaattaaatt	aagaaggaca	atgggtgccc	aatccttata	caaccaactt	17340
ccacaagaaa	gtcaagtccg	agacaacaaa	aaaacaagca	aaggaaattt	tttaatttga	17400
gttgtcttgt	ttgctgcata	atttatgcag	taaaacacta	cacataaccc	ttttagcagt	17460
agagcaatgg	ttgaccgtgt	gcttagcttc	ttttatttta	tttttttatac	agcaaagaat	17520
aaataaaata	aatgagaca	cttcagggat	gtttcaacaa	gcttggatcc	tcgaagagaa	17580
gggttaataa	cacctttttt	taacattttt	aacacaaatt	ttagttattt	aaaaatttat	17640
taaaaaattt	aaaataagaa	gaggaactct	ttaaataaat	ctaacttaca	aaatttatga	17700
tttttaataa	gttttcacca	ataaaaaatg	tcataaaaaat	atgttaaaaa	gtatattatc	17760
aatattctct	ttatgataaa	taaaaagaaa	aaaaaaataa	aagttaagtg	aaaatgagat	17820
tgaagtgaact	ttaggtgtgt	ataaatatat	caaccccgcc	aacaatttat	ttaatccaaa	17880
tatattgaag	tatattatc	catagccttt	atttatttat	atatttatta	tataaaagct	17940

-continued

---

ttatttggtc	taggttggtc	atgaaatatt	tttttggttt	tatctccggt	gtaagaaat	18000
catgtgcttt	gtgtgccac	tcactattgc	agctttttca	tgcattgggc	agattgacgg	18060
ttgattgtat	ttttggtttt	tatgggtttg	tgttatgact	taagtcttca	tctctttatc	18120
tcttcatcag	gtttgatggt	tacctaatat	ggtecatggg	tacatgcatg	gttaaattag	18180
gtggccaact	ttgttgtgaa	cgatagaatt	ttttttatat	taagtaaact	atttttatat	18240
tatgaaataa	taataaaaaa	aatattttat	cattattaac	aaaatcatat	tagttaattt	18300
gttaactcta	taataaaaga	aactactgaa	cattcacatt	acatggtaac	atctttccac	18360
cctttcattt	gttttttggt	tgatgaactt	ttttcttggt	taaatttatt	tcccttcttt	18420
taaatttggg	atacattatc	atcatatata	aactaaaata	ctaaaaacag	gattacacaa	18480
atgataaata	ataacacaaa	tatttataaa	tctagctgca	atatatttaa	actagctata	18540
tcgatattgt	aaaataaaac	tagctgcatt	gatactgata	aaaaaatatc	atgtgctttc	18600
tggactgatg	atgcagtata	cttttgacat	tgccctttatt	ttatttttca	gaaaagcttt	18660
cttagttctg	ggttcttcat	tatttgtttc	ccatctccat	tgtgaattga	atcatttget	18720
tcgtgtcaca	aatacaattt	agntaggtag	atgcattggg	cagattcacg	gtttattatg	18780
tcatgactta	agttcatggg	agtacattac	ctgccacgca	tgcattatat	tggttagatt	18840
tgataggcaa	atttggttgt	caacaatata	aatataaata	atgtttttat	attacgaaat	18900
aacagtgatc	aaaacaaaca	gttttatctt	tattaacaag	attttgtttt	tgtttgatga	18960
cgttttttaa	tgtttacgct	tcccccttc	ttttgaattt	agaacacttt	atcatcataa	19020
aatcaaatc	taaaaaaatt	acatatttca	taaataataa	cacaaatatt	tttaaaaaat	19080
ctgaaataat	aatgaacaat	attacatatt	atcacgaaaa	ttcattaata	aaaatattat	19140
ataaataaaa	tgtaaatagta	gttatatgta	ggaaaaaagt	actgcacgca	taatataatc	19200
aaaaagatta	aaatgaacta	ttataaataa	taaacactaa	ttaatggtga	atcatatcaa	19260
aataatgaaa	aagtaataaa	aatttgtaat	taacttctat	atgtattaca	cacacaaata	19320
ataaataata	gtaaaaaaaa	ttatgataaa	tatttaccat	ctcataagat	atttaaaata	19380
atgataaaaa	tatagattat	tttttatgca	actagctagc	caaaaagaga	acacgggtat	19440
atataaaaag	agtaccttta	aatttactg	tacttcttt	attcctgacg	tttttatatc	19500
aagtgacat	acgtgaagat	ttaattatc	agtctaaata	tttcattagc	acttaatact	19560
tttctgtttt	attcctatcc	tataagtagt	cccattctc	ccaacattgc	ttattcacac	19620
aactaactaa	gaaagtcttc	catagcccc	caagcggccc	atggcctcct	ccgaggacgt	19680
catcaaggag	ttcatgctc	tcaagtgctg	catggagggc	tccgtgaacg	gccacgagtt	19740
cgagatcgag	ggcgagggcg	agggccgccc	ctacgagggc	accagaccg	ccaagctgaa	19800
ggtgaccaag	ggcgcccccc	tgcccttcgc	ctgggacatc	ctgtcccccc	agttccagta	19860
cggtccaag	gtgtacgtga	agcaccctgc	cgacatcccc	gactacaaga	agctgtcctt	19920
ccccgagggc	ttcaagtggg	agcgcgtgat	gaacttcgag	gacggcgggc	tggtgaccgt	19980
gaccaggac	tcctcctgc	aggacggctc	cttcatctac	aaggtgaagt	tcacggcgt	20040
gaactcccc	tccgacggcc	ccgtaatgca	gaagaagact	atgggctggg	aggcctccac	20100
cgagcgcctg	tacccccgcg	acggcgtgct	gaagggcgag	atccacaagg	ccctgaaact	20160
gaaggacggc	ggccactacc	tggtggagtt	caagtccatc	tacatggcca	agaagcccgt	20220
gcagctgccc	ggctactact	acgtggactc	caagctggac	atcacctccc	acaacgagga	20280

-continued

---

```

ctacaccatc gtggagcagt acgagcgcgc cgagggccgc caccacctgt tctgtagcg 20340
gccggccgcg acacaagtgt gagagtacta aataaatgct ttggtgtac gaaatcatta 20400
cactaaataa aataatcaaa gcttatatat gccttcgct aaggccgaat gcaaagaaat 20460
tggttcttctc tcgttatctt ttgccacttt tactagtagc tattaattac tacttaatca 20520
tctttgttta cggctcatta tatccgctga cgg 20553

```

```

<210> SEQ ID NO 72
<211> LENGTH: 586
<212> TYPE: DNA
<213> ORGANISM: Brassica napus

```

```

<400> SEQUENCE: 72

```

```

ccaatttatt atcatctatt tctgacatt ttaatccatc cacctatgtc aaaaacttat 60
agaaaatgtc aacttccaaa caaaacataa ttgaacttcg caaataaatt cttaataata 120
ttaaaaaatg ttacttaatt atttcttcaa ccccatcttc cgcgcgtagc ggggacaaag 180
actctagtta aatatagaag tttccgattc tcatcgata aaacggtgac tttggcgggc 240
tttcatgtgt aacaaattgg tttacaaaac cactgcctag tcgttttagtg tagaatcagc 300
gcatggaact cggattggag cgtgactttc acgtrccgga ggccccaccac cwcagcgggc 360
gttacgctct aagaatctcg cccacggttt tcttcatctc cccccgccca agtgtctccc 420
tcggtcgcca cttctcatca tgttacaggg accataaaaa tggcgatatt cttcagcccc 480
gggtataaat acacacatga tctgtggtg ggttcttcca caagttacat ctctctctgg 540
tttttgatt gcaagtgttt gtattttttg cctccgagag aaaatc 586

```

```

<210> SEQ ID NO 73
<211> LENGTH: 1924
<212> TYPE: DNA
<213> ORGANISM: Brassica napus
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: (859)..(859)
<223> OTHER INFORMATION: n is a, c, g, or t

```

```

<400> SEQUENCE: 73

```

```

ctatagatgg gatgaagctg ctctcgacaa atctgataaa actaagaag gttagtaatc 60
aatTTTTtaca aatcataga ttattTTTT cattgaatta tttttatgct ataccaagaa 120
ttgtatttta gtatttgttt taactacata taatagaatt aactacatat aaattaacta 180
aacttaaaat aaaaatagat ttgtttcctg aaattatttt aagaatatat atgtatatat 240
ctaaaatctt agacttagat agatttttct atctatctat tttggttact taaaataaat 300
aaatttgat aaataattgt atagttatca aaaattaaaa ctaatttttt taaagttgtt 360
gatataataa atactaaaga tttaacgatt aagtatttat ttaagtatag aattttgttt 420
tttttttaag tttagttatg aagttgtaa ttatatataa acaaaacaat atttcgaaat 480
tttattatca tattcgaata tatttttttt agtgatgatg tatgaattat taccataatt 540
tgaaagttta ctaaaaaata tatcaaatg aattgtaata tatgagttat taccttaacc 600
aaaattataa ataacatta aatataatta tatatgcat atttagccat acaatgtgtc 660
atcaatatta atagtcatgt caatattaca taatgccaat attatgctac ttaaacccca 720
aatcccctaa ctcccgttaa gtgcccacat tcataaatat acttattoga caaaataaaa 780
aactttaaaa tatttactaa tccgaccatg cacaagcatc cattccctat tccattgcca 840
cgggataaca atgcaacna ctctcaaaa aaagaaaaat tcaagctctt ttgcaaaaaa 900

```

-continued

```

aaataaaata attttaacac ctaaaatttt ttgtttccaa acttctacag ggaacacaca   960
taaaagaaaa agaggacgtc cactcggatc acgcaacaaa ccaaaagggtg tgtcatgact   1020
cctaagatat aatatttcoct tattcaaaat cataccattt taaattatga atgtatttcg   1080
tagtccacca gatatgtaat ccaccagcgt tcaaaccaaa gttttatgat tgtaagttta   1140
agtgaattat aataatatat tcttcacggt atcttttcat aactaattga gttatcaaac   1200
ttgatcgcac atgtggcttt gatagggtgtg actttttatgg tatacaattc tttcaaccta   1260
aaaacattat tgttccctcaa tatcttacat tagcttgac tgcaacaaaa tattttctca   1320
tctgttttct tcctttaaac caatttatta tcacttattt cctgacattt taatccatcc   1380
acctatgtca aaaacttata gaaaatgtca acttccaaac aaaacataat tgaacttcgc   1440
aaataaattc ttaataatat taaaaaatgt tacttaatta tttcttcaac cccattttcc   1500
gcgcgtagcg cggacaaaaga ctctagttaa atatagaagt ttccgattct catcgtataa   1560
aacggtgact ttggcgggct ttcatgtgta acaaattggt ttaacaaacc actgcctagt   1620
cgtttagtgt agaatcagcg catggaactc cgattggagc gtgactttca cgtgcccggag   1680
gcccaccacc acagcggggcg ttacgctcta agaatctcgc ccacggtttt cttcactcgc   1740
cccccgccaa gtgtcttcoct cgttcgccac ttctcaccaa gttacaggaa ccctaaaaat   1800
ggccttttct cagcccccgc tataatacac acatgatcct atagtggggt ttccacaag   1860
ttacatctcc ttctggattg tacatttcaa gtgtttgtgt tttttctgcc tctgagagaa   1920
aatc                                                                 1924

```

```

<210> SEQ ID NO 74
<211> LENGTH: 444
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: ODP1 consensus sequence
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: 2, 3, 4, 5, 6, 7, 9, 10, 11, 12, 13, 14, 16, 17, 18, 19,
20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33,
34, 35, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48,
50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63,
64, 65
<223> OTHER INFORMATION: Xaa = any amino acid
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 82,
103, 104, 105, 106, 107, 108, 109, 110, 113, 115, 124, 127, 130,
131, 142, 143, 144, 145, 150, 152, 154, 155, 157, 158, 161,
162, 164, 165, 175, 183, 206, 208, 216, 217, 224
<223> OTHER INFORMATION: Xaa = any amino acid
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: 228, 236, 246, 248, 249, 250, 251, 252, 253, 254, 255,
256, 257, 258, 259, 260, 261, 262, 263, 264, 265, 266, 267, 268,
269, 270, 271, 272, 273, 274, 275, 276, 277, 278, 279, 280,
281, 282, 283, 284, 285, 287, 288, 289, 290, 291, 292
<223> OTHER INFORMATION: Xaa = any amino acid
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: 293, 294, 295, 296, 297, 298, 299, 300, 302, 303, 304,
305, 306, 307, 308, 309, 310, 311, 312, 313, 314, 315, 316, 317,
318, 319, 320, 321, 322, 323, 324, 325, 326, 327, 328, 329,
331, 333, 334, 335, 336, 337, 338, 339, 340, 341, 342
<223> OTHER INFORMATION: Xaa = any amino acid
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: 343, 344, 345, 346, 347, 348, 349, 350, 351, 352, 353,
354, 355, 356, 357, 358, 359, 360, 361, 362, 363, 364, 365, 366,
367, 368, 370, 371, 372, 374, 375, 376, 378, 379, 380, 381,
382, 383, 384, 385, 386, 387, 388, 389, 390, 391, 392

```

-continued

---

```

<223> OTHER INFORMATION: Xaa = any amino acid
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: 393, 394, 395, 396, 397, 398, 399, 400, 385, 386, 387,
388, 389, 390, 391, 392, 393, 394, 395, 396, 397, 398, 399, 400,
401, 402, 403, 404, 405, 406, 407, 408, 409, 410, 411, 412, 413,
414, 415, 416, 417, 418, 419, 420, 421, 422, 423, 424, 425, 426,
427
<223> OTHER INFORMATION: Xaa = any amino acid
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: 428, 429, 430, 431, 432, 433, 434, 435, 436, 437, 439,
440, 441, 442, 443, 444
<223> OTHER INFORMATION: Xaa = any amino acid

<400> SEQUENCE: 74

Met Xaa Xaa Xaa Xaa Xaa Xaa Ser Xaa Xaa Xaa Xaa Xaa Xaa Ser Xaa
1          5          10          15

Xaa Xaa
20          25          30

Xaa Xaa Xaa Arg Xaa Xaa
35          40          45

Arg Xaa Xaa
50          55          60

Xaa Arg Ser
65          70          75          80

Ser Xaa Tyr Arg Gly Val Thr Arg His Arg Trp Thr Gly Arg Phe Glu
85          90          95

Ala His Leu Trp Asp Lys Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Lys Lys
100         105         110

Xaa Gly Xaa Gln Val Tyr Leu Gly Ala Tyr Asp Xaa Glu Glu Xaa Ala
115         120         125

Ala Xaa Xaa Tyr Asp Leu Ala Ala Leu Lys Tyr Trp Gly Xaa Xaa Xaa
130         135         140

Xaa Leu Asn Phe Pro Xaa Glu Xaa Tyr Xaa Xaa Glu Xaa Xaa Glu Met
145         150         155         160

Xaa Xaa Val Xaa Xaa Glu Glu Tyr Leu Ala Ser Leu Arg Arg Xaa Ser
165         170         175

Ser Gly Phe Ser Arg Gly Xaa Ser Lys Tyr Arg Gly Val Ala Arg His
180         185         190

His His Asn Gly Arg Trp Glu Ala Arg Ile Gly Arg Val Xaa Gly Xaa
195         200         205

Lys Tyr Leu Tyr Leu Gly Thr Xaa Xaa Thr Gln Glu Glu Ala Ala Xaa
210         215         220

Ala Tyr Asp Xaa Ala Ala Ile Glu Tyr Arg Gly Xaa Asn Ala Val Thr
225         230         235         240

Asn Phe Asp Ile Ser Xaa Tyr Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
245         250         255

Xaa Xaa
260         265         270

Xaa Glu Xaa Xaa
275         280         285

Xaa Pro Xaa Xaa Xaa
290         295         300

Xaa Xaa
305         310         315         320

Xaa Leu Xaa Trp Xaa Xaa Xaa Xaa
325         330         335

```

-continued

---

Xaa  
 340 345 350

Xaa  
 355 360 365

Phe Xaa Xaa Xaa Ile Xaa Xaa Xaa Phe Xaa Xaa Xaa Xaa Xaa Xaa Xaa  
 370 375 380

Xaa  
 385 390 395 400

Xaa  
 405 410 415

Xaa  
 420 425 430

Xaa Xaa Xaa Xaa Xaa Cys Xaa Xaa Xaa Xaa Xaa Xaa  
 435 440

---

What is claimed is:

1. A recombinant DNA construct comprising a polynucleotide encoding an ODP1 polypeptide operably linked to a sucrose synthase 2 promoter (SUS2) wherein the SUS2 promoter comprises a nucleotide sequence having at least 95% sequence identity to SEQ ID NO:43, wherein said nucleotide sequence has seed specific promoter activity and wherein the amino acid sequence of said ODP1 polypeptide has at least 90% sequence identity to SEQ ID NO:39 and comprises two APETALA2 (AP2) domains and wherein expression of said ODP1 polypeptide increases oil content in the seeds of a cruciferous oilseed plant while maintaining normal germination.

2. The recombinant DNA construct of claim 1 wherein the amino acid sequence of said ODP1 polypeptide has at least 95% sequence identity to SEQ ID NO:39.

3. The recombinant DNA construct of claim 1 wherein the amino acid sequence of said ODP1 polypeptide comprises SEQ ID NO:39.

4. The recombinant DNA construct of claim 1 wherein the sucrose synthase 2 promoter comprises the nucleotide sequence of SEQ ID NO:43.

5. The recombinant DNA construct of claim 2, wherein the sucrose synthase 2 promoter comprises the nucleotide sequence of SEQ ID NO:43.

6. The recombinant DNA construct of claim 3, wherein the sucrose synthase 2 promoter comprises the nucleotide sequence of SEQ ID NO:43.

7. The recombinant DNA construct of claim 1 wherein the oilseed plant is canola or *Arabidopsis*.

8. A transgenic cruciferous oilseed plant comprising in its genome the recombinant DNA construct of claim 1.

9. The transgenic cruciferous oilseed plant of claim 8 wherein the cruciferous oilseed plant is selected from the group consisting of canola and *Arabidopsis*.

10. A transgenic seed obtained from the plant of claim 8, wherein said seed comprises in its genome said recombinant DNA construct.

11. The transgenic cruciferous oilseed plant of claim 8, wherein the amino acid sequence of said ODP1 polypeptide comprises SEQ ID NO:39.

12. The transgenic cruciferous oilseed plant of claim 8, wherein the sucrose synthase 2 promoter comprises the nucleotide sequence of SEQ ID NO:43 and wherein the

amino acid sequence of said ODP1 polypeptide has at least 95% sequence identity to SEQ ID NO:39.

13. The transgenic cruciferous oilseed plant of claim 8, wherein the sucrose synthase 2 promoter comprises the nucleotide sequence of SEQ ID NO:43 and wherein the amino acid sequence of said ODP1 polypeptide comprises SEQ ID NO:39.

14. A transgenic seed obtained from the plant of claim 8, wherein said seed comprises in its genome said recombinant DNA construct and wherein the amino acid sequence of said ODP1 polypeptide comprises SEQ ID NO:39.

15. A transgenic seed obtained from the plant of claim 8, wherein said seed comprises in its genome said recombinant DNA construct and wherein the sucrose synthase 2 promoter comprises the nucleotide sequence of SEQ ID NO:43 and wherein the amino acid sequence of ODP1 polypeptide has at least 95% sequence identity to SEQ ID NO:39.

16. A transgenic seed obtained from the plant of claim 8, wherein said seed comprises in its genome said recombinant DNA construct and wherein the sucrose synthase 2 promoter comprises the nucleotide sequence of SEQ ID NO:43 and wherein the amino acid sequence of said ODP1 polypeptide comprises SEQ ID NO:39.

17. A method for producing a transgenic cruciferous oilseed plant comprising transforming a cruciferous oilseed plant cell with the recombinant DNA construct of claim 1 and regenerating a transgenic cruciferous oilseed plant from the transformed cruciferous oilseed plant cell, wherein the transgenic cruciferous oilseed plant comprises in its genome said recombinant DNA construct.

18. The method of claim 17 wherein the cruciferous oilseed plant is selected from the group consisting of canola and *Arabidopsis*.

19. A method for increasing oil content in seeds of a transgenic cruciferous oilseed plant while maintaining normal germination, said method comprising:

- (a) transforming a cruciferous oilseed plant cell with a recombinant DNA construct comprising a polynucleotide encoding an ODP1 polypeptide, wherein the amino acid sequence of said ODP1 polypeptide has at least 90% sequence identity to SEQ ID NO:39 and comprises two APETALA2 (AP2) domains, said polynucleotide being operably linked to a promoter having a nucleotide sequence at least 95% identical to SEQ ID NO: 43, wherein said nucleotide sequence has seed specific promoter activity;

- (b) regenerating a transgenic cruciferous oilseed plant from the transformed cell of step (a), wherein said plant comprises the recombinant DNA construct;
- (c) obtaining a transgenic progeny plant derived from the transgenic cruciferous oilseed plant of step (b), wherein the transgenic progeny plant comprises in its genome the recombinant DNA construct; 5
- (d) assaying the transgenic progeny plant obtained from step (c) for oil level and germination; and
- (e) selecting those transgenic progeny plants having seeds comprising said recombinant DNA construct and having an increased level of oil and normal germination when compared to seeds obtained from a control cruciferous oilseed plant, wherein said control plant does not comprise the recombinant DNA construct. 15

20. The method of claim 19 wherein the amino acid sequence of the ODP1 polypeptide comprises the sequence of SEQ ID NO:39.

21. The method of claim 19 wherein the promoter comprises SEQ ID NO:43. 20

22. The method of claim 21 wherein the ODP1 polypeptide comprises at least 95% sequence identity to SEQ ID NO: 39.

23. The method of claim 19 wherein the cruciferous oilseed plant is canola or *Arabidopsis*.

\* \* \* \* \*

25